NtdB:
A kanosamine-6-phosphate phosphatase

A Thesis Submitted to the
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in the Department of Chemistry
University of Saskatchewan
Saskatoon

By
Julie Boisvert-Martel

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Chemistry
University of Saskatchewan
Saskatoon, Saskatchewan (S7N 5C9)

ABSTRACT

NtdB is an enzyme encoded within the *ntd* operon in *Bacillus subtilis*. This operon is reported to contain a complete set of genes necessary for the biosynthesis of 3,3'-neotrehalosadiamine (NTD), a compound composed of two kanosamine subunits linked together by a 1,1'-(α,β)-linkage. Both NTD and kanosamine have reported antibiotic properties. The function of NtdB has been a matter of speculation, but has never been investigated *in vitro*. Using a phosphate assay and an array of substrates, NtdB was determined to be a phosphatase, specific to kanosamine-6-phosphate (K6P) (k_{cat} = 32 ± 1 s^{-1}, K_m = 93 ± 7 μM). Site-directed mutagenesis of amino acid residues in the core and the cap domains of the enzyme identified residues important for the catalytic reaction and substrate specificity. These mutations confirmed the presence of four motifs, characteristic of members of the haloacid dehalogenase (HAD) superfamily, and allowed identification of the substrate binding site of the enzyme. KabB, a homologue of NtdB from *Bacillus cereus*, showed notably lower activity with K6P than NtdB. This research defines the role of NtdB as a specific K6P phosphatase and challenges the previously reported NTD biosynthesis pathway by demonstrating a novel pathway for the production of the antibiotic kanosamine.

ACKNOWLEDGEMENTS

This thesis would have remained a dream without the immense help and support of my supervisor, Dr. David R. J. Palmer. I cannot find the words to express my appreciation for having given me such an interesting and challenging research project.

I consider it an honor to have worked with my lab manager, Natasha Vetter. Her patience, dedication, and expertise were vastly appreciated and I do not know what I would have done without her.

I am indebted to my professor and co-supervisor, Dr. David A. R. Sanders. He was the first contact I had with the University of Saskatchewan and his sense of humor, and expertise, have contributed to make this experience memorable.

Thank you to the members of the Palmer Lab (Yulia Skovpen, Dr. Shazia Anjum, Dr. Rajendra Jagdhane, Hari Babu Aamudalapalli, Aarti Bhagwat, Bernardo Jung), and of the Sanders lab (Dr. Karin E. van Straaten, Dr. Sean Dalrymple, Jijin Raj Ayanath Kuttiyatveetil, Carla Protsko, Drew Bertwistle, Dr. Mary Yang, Siddharth Tiwari, Cuylar Conly) for their constant support and suggestions.

It is with gratitude that I acknowledge the support and help of the Saskatchewan Structural Science Center staff, more specifically Jason Maley and Ken Thoms. Their teachings and expertise were much appreciated.

Finally, I owe a huge thank you to my family. This thesis would not have been possible without the love and support of my mother and my three kids.

DEDICATION

Cette thèse est dédiée à Mélanie, Catherine et Alexandra.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

LIST OF ABBREVIATIONS

AMMP	2-Amino-6-mercapto-7-methylpurine
Bis-Tris	Bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane
CD	Circular dichroism
HAD	Haloacid dehalogenase
G6P	D-Glucose-6-phosphate
GDP	Guanosine diphosphate
GlcAm6P	Glucosamine-6-phosphate
K1P	Kanosamine-1-phosphate
K6P	Kanosamine-6-phosphate
LB	Luria-Bertani
NAD^+	Nicotinamide adenine dinucleotide
NTD	3,3'-neotrehalosadiamine
PCR	Polymerase chain reaction
PLP	Pyridoxal phosphate
PMP	Pyridoxamine phosphate
PTPase	Protein tyrosine phosphatase
R1P	Ribose-1-phosphate
S1	Strand one
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
T_m	Melting temperature
TPP	Trehalose-6-phosphate phosphatase
UDP	Uridine diphosphate
WT	Wild type

1 INTRODUCTION

1.1 Discovering new antibiotics, such as 3,3'-neotrehalosadiamine

In 1986, in the course of a screening for aminoglycoside antibiotics, a new aminosugar antibiotic produced by *Bacillus sp.*, 3,3'-neotrehalosadiamine (NTD), was reported.[1] Using a paper disc assay, it was shown to inhibit the growth of *Staphylococcus aureus* and *Klebsiella pneumoniae*.[2] During the last 60 years, *S. aureus* (both hospital and community acquired) has developed resistance not only to methicillin, but also to oxacillin (replacement to methicillin) and other beta-lactams.[3] Some methicillin-resistant strains of *S. aureus* (MRSA) are also becoming resistant to erythromycin, ciprofloxacin, gentamicin, neomycin, trimethoprim, and tetracycline.[4] Similarly, new strains of carbapenem-resistant *K. pneumonia* (CRKP) have started appearing in hospitals all over the world, with no efficient antibiotics to fight them.[5,6] Therefore, in parallel with research on the mechanisms of resistance of these bacteria, the investigation of new potential antibiotics, such as NTD, is essential (**Figure 1.1**).

NTD is composed of two kanosamine subunits, linked together by a α,β-1,1'-linkage (**Figure 1.1**).[2] This type of α,β-linkage is related to neotrehaloses and, to this day, NTD is the only known neotrehalose composed of two aminosugars. NTD was first isolated as part of a screening program to obtain novel aminoglycoside antibiotics.[2] It was reported to be produced naturally by *B. pumilis*[2] and *B. circulans*[7] as well as a rifampicin resistant mutant of *B. subtilis*.[8]

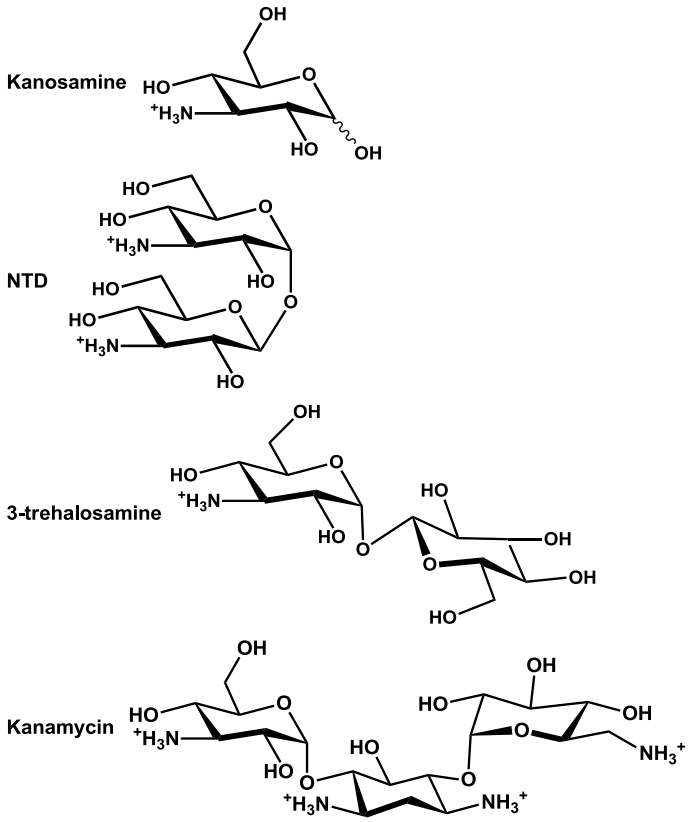


Figure 1.1: Structure of aminoglycoside antibiotics

It is not surprising that NTD exhibits antibiotic properties. Kanosamine has antibacterial and antifungal activity,[9, 10] and is reported to inhibit bacterial cell walls synthesis.[11] The kanosamine subunit is also found in many other antibiotics, such as kanamycin and α,α-3-trehalosamine (**Figure 1.1**).[12,13]

1.2 *ntd* operon

A Rif^r*rpoB5* mutation in *B. subtilis* has been shown to activate a dormant ability to produce NTD.[8] The genes involved in the production of this antibiotic were isolated using transposon mutagenesis, which consisted of the insertion of a vector which interrupted the gene activity. Three genes were identified, *ntdA*, *ntdB*, and *ntdC*, and their role in the biosynthesis of NTD was confirmed by cloning them into pUC18 and expressing them in a bacteria not affected by NTD, *Escherichia coli* (*E. coli*), under the

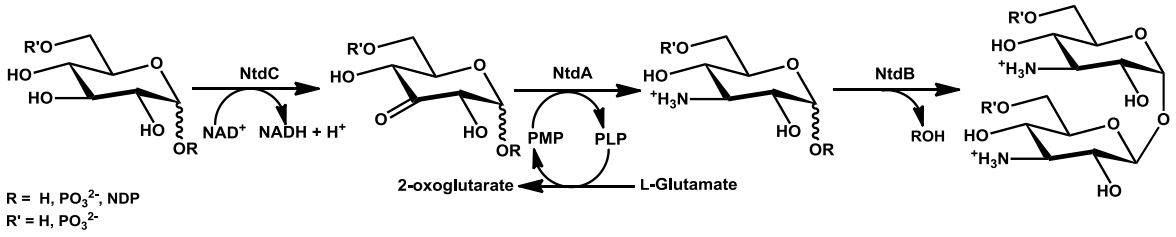
control of a *lac* promoter. The unusual production of NTD by *E. coli* was confirmed by ESI-MS analysis.[8]

Further investigation revealed that the *ntd* operon contains a total of 5 genes. The first is a transcriptional regulator, *ntdR*, which is encoded in reverse compared to the others genes of the operon. The mechanism obeyed by this regulator has yet to be understood, but preliminary studies seem to indicate that it may function as an autoinducer.[14] This gene is followed by *ntdA*, *ntdB*, and *ntdC*. There is a sequence overlap of *ntdA* and *ntdB*, such that the beginning of *ntdB* is encoded in the end of *ntdA* by 29 base pairs. The last gene in the operon is *glcP*, which encodes for a glucose/mannose:H^+ symport permease.[8, 14]

1.3 End product of the operon

1.3.1 First hypothesis: NTD

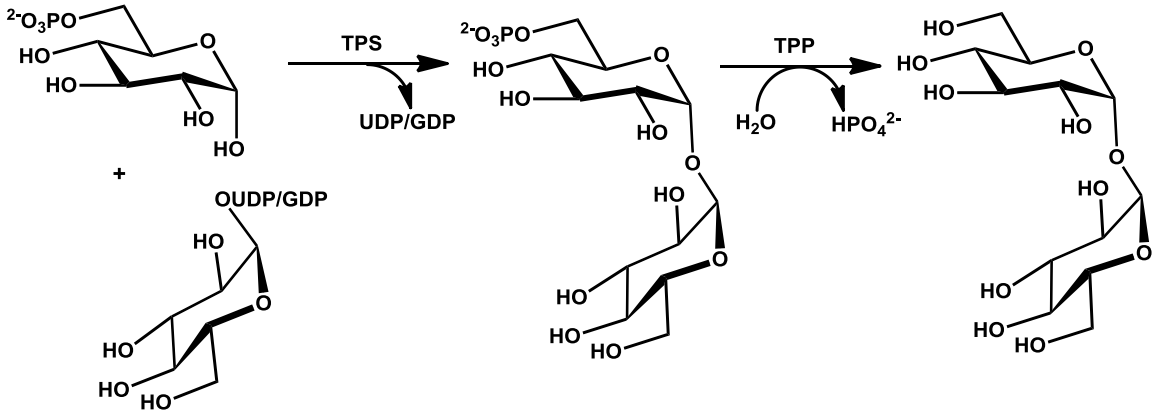
To understand the production of NTD by the *ntd* operon, study of the three genes, *ntdABC*, was undertaken. According to their amino acid sequence, NtdA was suggested to be a pyridoxal phosphate (PLP)-dependent aminotransferase, NtdB a hydrolase, and NtdC a nicotinamide adenine dinucleotide (NAD^+) dependent dehydrogenase.[8] Based on the proposed functions of the three enzymes encoded by the *ntdABC* operon, an NTD biosynthetic pathway was initially proposed (**Scheme 1.1**). This hypothesis was tested by Palmer *et al*.[15, 16]



Scheme 1.1: The biosynthetic pathway of *ntdABC* operon, producing NTD as a final product.

The kinetics of NtdC were evaluated and showed that glucose-6-phosphate (G6P) was a good substrate for this enzyme, with a k_{cat} of 4.0 ± 0.1 s^{-1} and a K_m of 13 ± 1 μM.[15] NtdA was characterized as a Glu-dependent aminotransferase[17], and its crystal structure was solved in its three main forms: with its cofactor, PLP, with pyridoxamine phosphate (PMP), and with the proposed product, kanosamine-6-phosphate (K6P).[16,18] It was also shown to be specific for α-D-K6P in the reverse direction of the reaction shown in **Scheme 1.1**.[18] The proposed conversion of G6P and K6P by NtdC and NtdA based on **Scheme 1.1**, was confirmed by HPLC, and a short lived intermediate was observed by UV spectrophotometry at 310 nm.[17]

The assignment of NtdB in **Scheme 1.1** presumed that the enzyme was responsible for the α,β-glycosidic bond found in NTD. Usually, such a bond is formed with a glycosyltransferase and a nucleotidyl sugar, as can be seen in **Scheme 1.2** in which trehalose-6-phosphate synthase (TPS) uses G6P and uridine diphosphate (UDP)- or guanosine diphosphate (GDP)-glucose to form trehalose-6-phosphate.[19] Dephosphorylation occurs in the second step, by the action of trehalose-6-phosphate phosphatase (TPP). This pathway is not consistent with the results and observations found for the NtdA and NtdC enzymes. Different hypotheses therefore arise: NtdB is a unique kind of enzyme, catalyzing a novel type of reaction; the *ntdABC* operon is incomplete, some genes such as a glycosyltransferase are missing; the *ntdABC* operon has a different end product than NTD. This last hypothesis is further discussed in the next section.



Scheme 1.2: The catalytic pathway of trehalose using G6P and UDP/GDP-glucose as initial substrates.

1.3.2 Second hypothesis: Kanosamine

Since 1996, it was known that *B. cereus* produces the antibiotic kanosamine.[9] Thirteen years later, the *kab* operon from *B. cereus* was identified as a potential operon containing the genes necessary for the production of kanosamine. A sequence alignment comparing KabA, KabB, and KabC with NtdA, NtdB, and NtdC was performed and revealed an identity of 57%, 55% and 57% respectively (**Figure 1.2-1.4**). This high sequence identity between both sets of proteins suggests that the *ntdABC* operon may encode for kanosamine production instead of NTD.

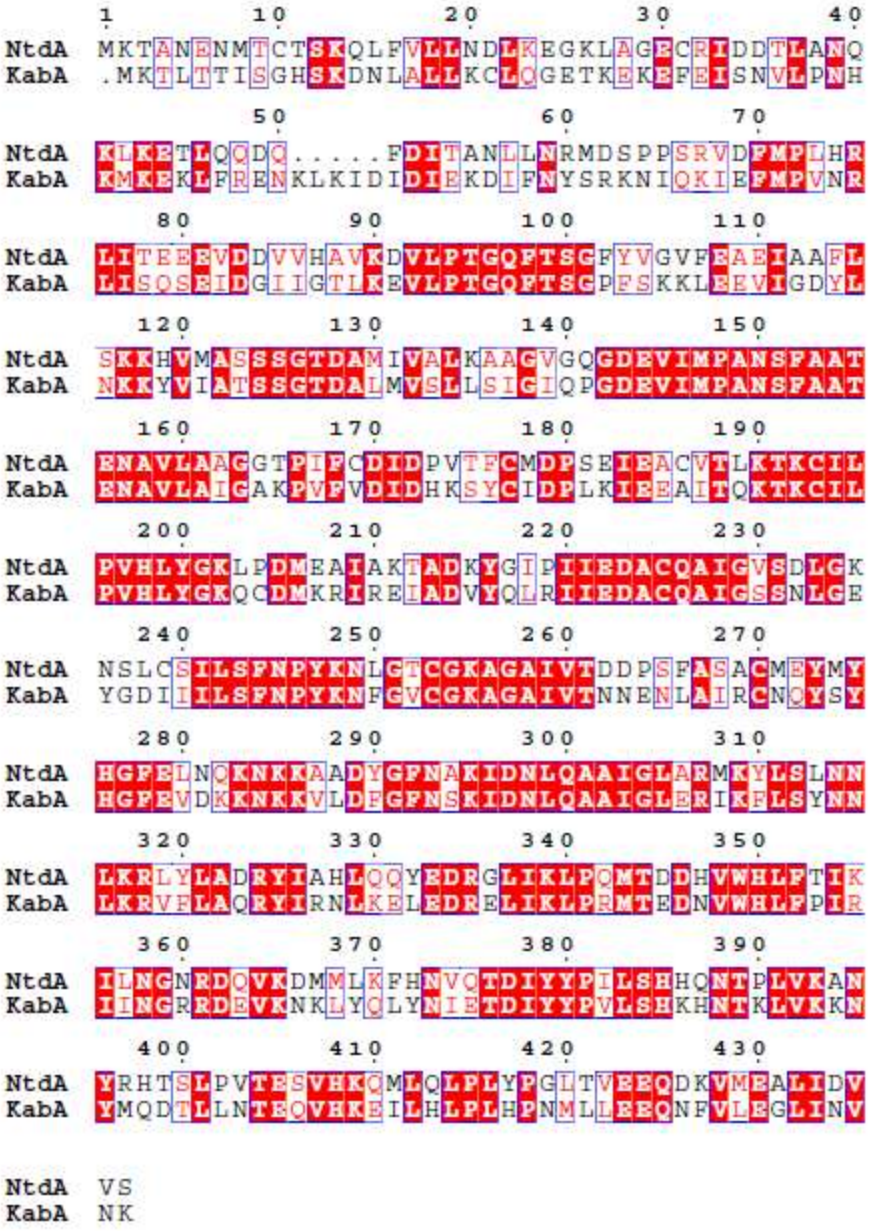


Figure 1.2: Sequence alignment of NtdA and KabA.
Generated with *ClustalW*[20], using the Blosum scoring matrix[21]: opening gap penality = 10; extending gap penality = 0.20; distance gap penality = 5. Image was created by *ESPript* V2.2[22]. The amino acids represented with white letters are identical, while red letters denote similar residues.

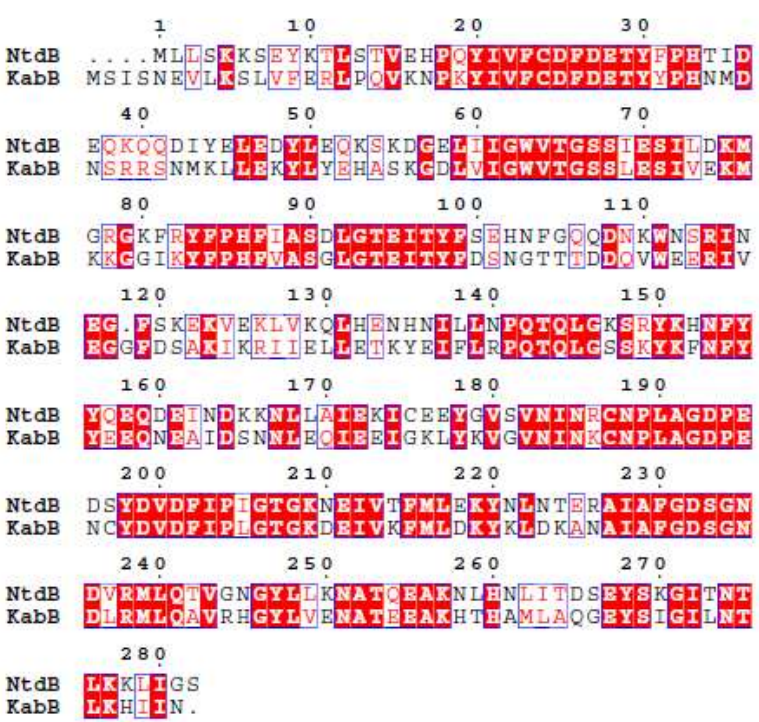


Figure 1.3: Sequence alignment of NtdB and KabB.
Generated with *ClustalW*[20], using the Blosum scoring matrix[21]: opening gap penality = 10; extending gap penality = 0.20; distance gap penality = 5. Image was created by *ESPript* V2.2[22]. The amino acids represented with white letters are identical, while red letters denote similar residues.

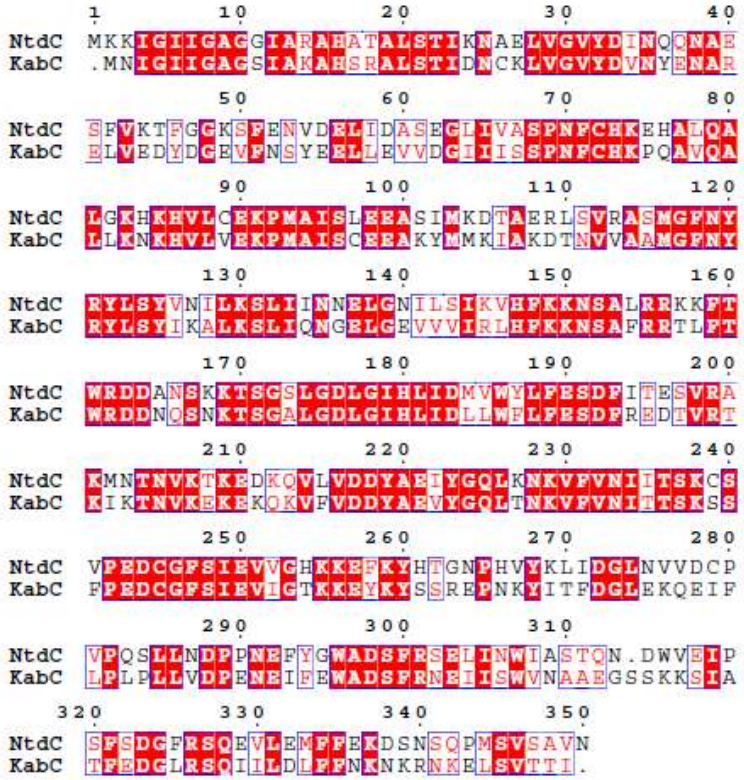
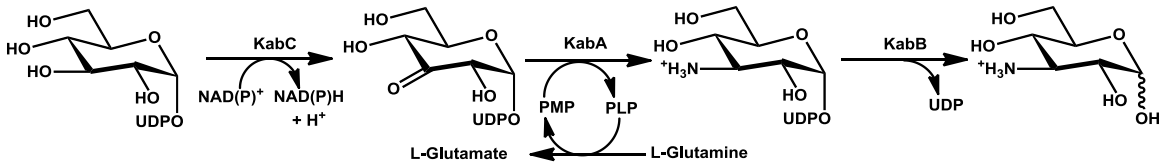
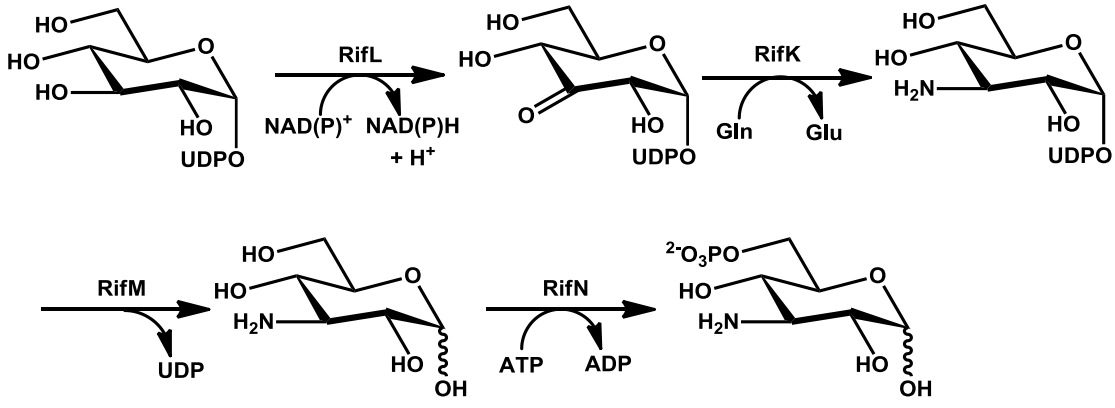


Figure 1.4: Sequence alignment of NtdC and KabC.
Generated with *ClustalW*[20], using the Blosum scoring matrix[21]: opening gap penality = 10; extending gap penality = 0.20; distance gap penality = 5. Image was created by *ESPript* V2.2[22]. The amino acids represented with white letters are identical, while red letters denote similar residues.



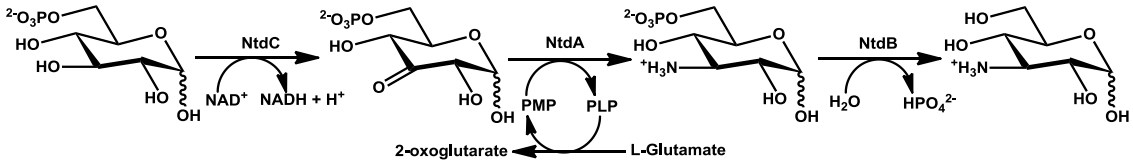
Scheme 1.3: The proposed biosynthetic pathway of kanosamine by the *kabABC* operon from *B. cereus*.[23]

Kevany *et al*. proposed a scheme for the *kab* operon (**scheme 1.3**),[23] based on the first steps of the proposed aminoshikimate pathway, which produces 3-amino-5-hydroxybenzoic acid (AHBA). Using electrospray mass spectrometry, Guo and Frost detected kanosamine and K6P as intermediates in this pathway, and it was suggested that they are the source of nitrogen for AHBA.[24] The biosynthesis of kanosamine is believed to be performed by three proteins, RifL (oxydoreductase), RifK (aminotransferase), and RifM (UDPase).[24] Overexpression of these three genes in *E. coli* was previously attempted and unsuccessful.[25] A fourth enzyme, RifN, phosphorylates kanosamine to K6P (**Scheme 1.4**).[26] A sequence alignment with the three first enzymes reveals a sequence identity of less than 16% with enzymes of the *kab* operon. This would suggest that the *rif* operon and the *kab* operons have different mechanisms for kanosamine biosynthesis.



Scheme 1.4: The proposed first steps of the aminoshikimate pathway.

Based on the high similarity between the *ntd* and *kab* operons, a biosynthetic pathway is proposed, whereby kanosamine is the end product (**Scheme 1.5**). It also takes into account previous data obtained on NtdA and NtdC. [15-17] This pathway would imply the presence of at least one more enzyme, found naturally in *B. subtilis* and *E. coli* but not in *B. cereus*, which would be responsible for the formation of the glycosidic bond. In this proposed scheme, NtdB would act as a phosphatase, removing the phosphate group of K6P.



Scheme 1.5: The proposed biosynthetic pathway of kanosamine by the *ntdABC* operon from *Bacillus sp*.

1.4 Hydrolysis of phosphate esters

A phosphate ester contains at least one carbon attached to the phosphate oxygen (**Figure 1.5**). Polyphosphate molecules contain more than one phosphate attached to one another, such as in ATP.

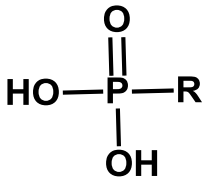


Figure 1.5: A phosphate monoester.
R = any group bonded to the oxygen atom through a carbon atom.

Hydrolysis of a phosphate ester is a thermodynamically favorable process which occurs via the attack from a water molecule on the phosphate. This process is slow due to the electrostatic repulsion between the negative charge of the oxygen surrounding the phosphate center and the oxygen of water.[27]

There are two hydrolysis mechanisms, dissociative and associative. The dissociative mechanism proceeds via a hydrated PO_3^- anion. The associative mechanism requires a transition state which involves the formation of pentacoordinated phosphate.[28] A study using methyl phosphate under different conditions, such as different protonation state of the substrate and two forms of nucleophile, water and hydroxide anion, revealed the difficulty in distinguishing between both mechanisms, since they have similar activation barriers.[28,29]

Uncatalyzed, this hydrolysis reaction may take thousands of years.[30,31] With the help of phosphatases, the hydrolysis of phosphate esters, occurs in seconds. Two mechanisms have been found to be used by enzymes, the direct transfer of the phosphoryl group to a metal coordinated water, and the formation of a phosphoenzyme intermediate.[32]

1.4.1 Metallophosphatases

Metallophosphatases use a one-step mechanism which involves one or two metal ions, such as magnesium, manganese, iron, or zinc, or a mixture of two, to help coordinate a hydroxide ion, which then acts as a nucleophile, attacking the phosphate group of the substrate (**Figure 1.6**).[33,34] Many of these phosphatases are found in the serine/threonine superfamily, more specifically the major families, phosphoprotein phosphatase (PPP) and the metal dependent protein phosphatase (PPM), as well as purple acid phosphatase.[32,33,35]

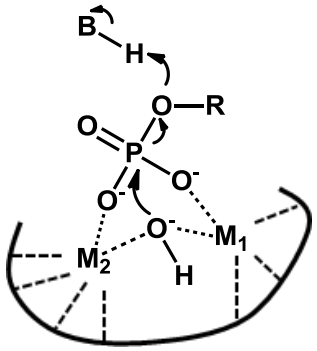


Figure 1.6: Coordination of a hydroxide ion by two metals.
The ion can act as a nucleophile, attacking the phosphate group. M_1 and M_2 represent two metal ions, such as zinc, iron, magnesium, or manganese, which are coordinated by residues in the enzyme.

1.4.2 Phosphoenzyme intermediate

The second mechanism involves a nucleophilic attack by an active site amino acid on the substrate leading to the formation of a phosphoenzyme intermediate, followed by a hydroxide nucleophilic attack on this intermediate, causing the release of the phosphate group from the enzyme (**Figure 1.7**). To coordinate the phosphate group, phosphatases use either amino acids within their active site or metal cofactors.

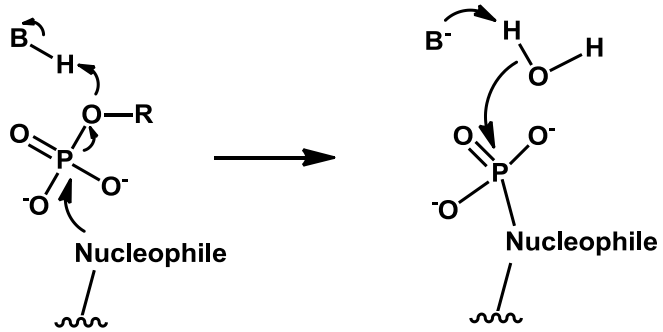


Figure 1.7: Two steps mechanism for phosphatases.
The attack of the nucleophile leads to the formation of a phosphoenzyme intermediate, which is attacked by a water molecule. The phosphate group is coordinated by either residues in the enzyme or metal cofactor.

Different residues were found to perform the first nucleophilic attack, such as cysteine, serine, histidine, or aspartate (**Table 1.1**). The same variability in the identity of the general acid/base is found. It was proposed that some protein tyrosine phosphatases (PTPases) do not use a general acid because the pK_a of the leaving group is low.[36, 37] For

phosphonatases in the haloacid dehalogenase (HAD) superfamily, a Schiff base mechanism is used to catalyze the reaction.[38]

Table 1.1: Comparison of different nucleophilic residues used by phosphatases.

Nucleophilic residue	Cysteine	Serine	Histidine	Aspartate
Major examples	PTPases[39]	Alkaline phosphatases[40,41]	Many acid phosphatases,[42] Glucose 6-phosphatase[43]	FCP/SCP[35] HAD[44]
Conserved active site	CxxxxxR	DSA/G	RHGxRxP	DxDxT/V
General acid/base	Aspartate or Glutamate[37] or None[36]	Zn^{2+} and Mg^{2+}	Histidine	Aspartate or Enamine intermediate[38]
Coordination of phosphate	Arginine	Arginine	Arginine	Mg^{2+}

Abbreviation: SCP, small C-terminal domain phosphatase; HAD, haloacid dehalogenase;

1.5 Assigning NtdB to a superfamily

A protein BLAST search[45] using the NtdB protein sequence as query, identifies this enzyme as belonging to the HAD superfamily. The HAD superfamily, characterized in 1994, was named after the haloacid dehalogenase and can be found in all organism domains.[46] Contrary to what the name indicates, this superfamily is primarily composed of enzymes catalyzing the transfer or removal of a phosphate group on a wide range of phosphorylated substrates.[44]

1.5.1 Structure of HAD phosphatases

Members of the HAD superfamily contain two distinct domains; the core, where catalysis occurs, and the cap, which regulates access to the active site by different means, such as limiting the solvent access or providing substrate specificity.[44] The cap domain can be further divided in three categories, C0, C1, and C2, based on its secondary structure. The crystal structure of NtdB from *B. subtilis* was solved in 2009[47] (PDB 3GYG, **figure 1.8**) and clearly shows the presence of these two distinct domains,

connected together via a hinge loop (P205-G208) and a partially disordered α-helix (N109-G118). The magnesium cofactor is bound in the core domain.

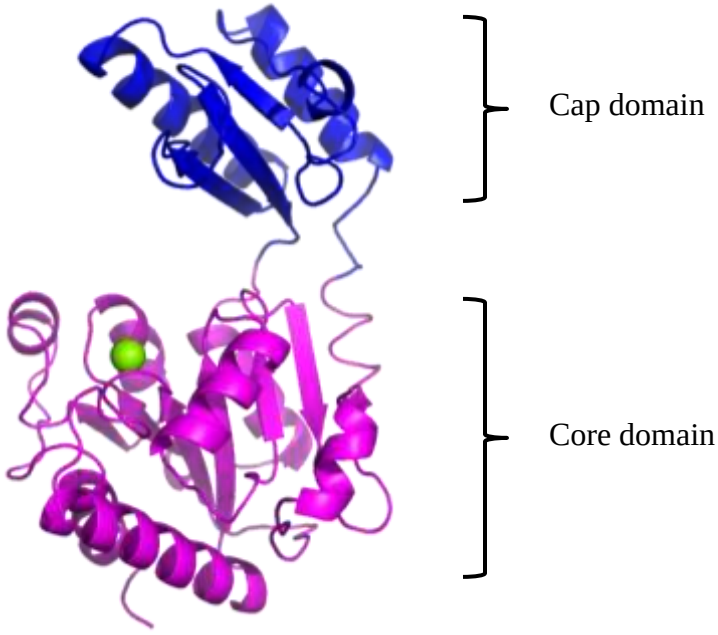


Figure 1.8: Crystal structure of NtdB.
The core domain is at the bottom and in magenta while the cap domain is at the top, in blue. Magnesium is indicated in green (PDB 3GYG).

The common overall structural arrangement favors the following catalytic cycle for members of the HAD superfamily possessing a C1 cap domain (**Figure 1.9**). In the absence of substrate, the enzyme is in the open conformation, allowing the substrate to bind to the active site located in the core. The magnesium cofactor helps to orient the phosphate group of the substrate. The enzyme adopts a closed-conformation, preventing any solvent from reaching the active site. A nucleophilic attack by aspartate on the substrate occurs within the active site, creating an aspartate-bound intermediate. Then, the enzyme reverts back to its open conformation, protonating and liberating the dephosphorylated substrate, and allowing solvent to access the active site. The water molecule is deprotonated by the general base and acts as a nucleophile, attacking the aspartate-bound intermediate. The inorganic phosphate is released and the enzyme returns to its original state.[44]

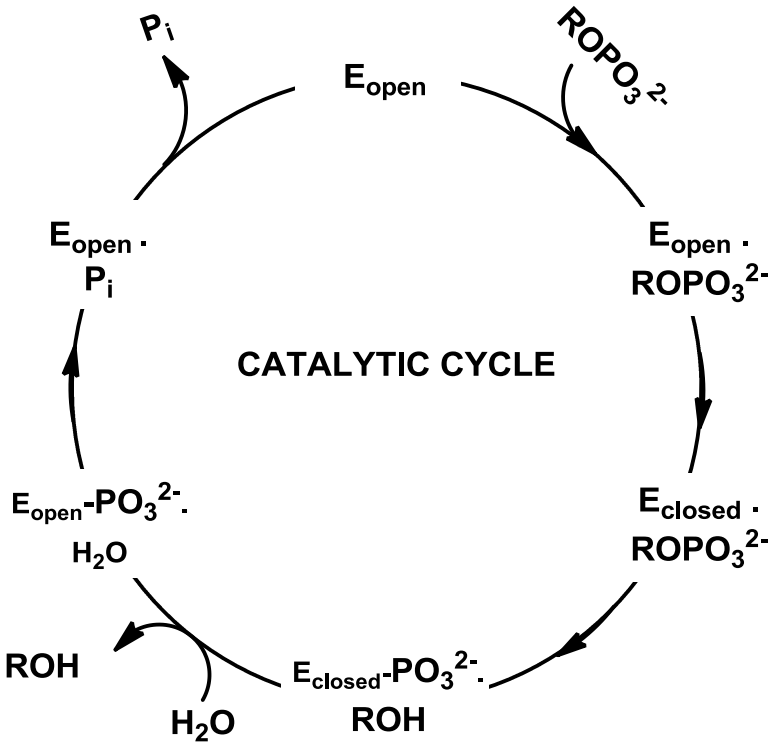


Figure 1.9: The catalytic cycle for members of the HAD superfamily possessing a C1 cap domain

The important structural feature emerging from this catalytic cycle is the alternations between open and closed conformations. This allows the solvent to be included or excluded depending on the step of the reaction. It was proposed that a similar cycle occurs for enzymes containing a C2 caps.[44]

1.5.2 Secondary structure of the core domain

The core domain of members of the HAD superfamily is characterized by the presence of a Rossmann-like fold. Typically, Rossmann folds are present when an enzyme binds nucleotides, such as NAD(P) or FMN. In general, its structure consists of an alternation of α-helices and β-strands with a central β-sheet, containing five parallel β-strands in a 54123 order (the numbering refers to the order of appearance in the amino acid sequence where S1 corresponds to Strand 1 and so on).[44,48] Additionally, there is a conserved motif, GxGxxG, where the phosphate group of the nucleotide binds.[49,50]

As it can be seen in **figure 1.10**, NtdB contains an alternation of β/α motifs, the central β-sheet containing five strands in its core (**Table 1.2**). The absence of the conserved sequence of three glycine residues in its amino acids sequence indicates that NtdB does not contain a typical Rossmann fold and is less likely to bind a nucleotide.

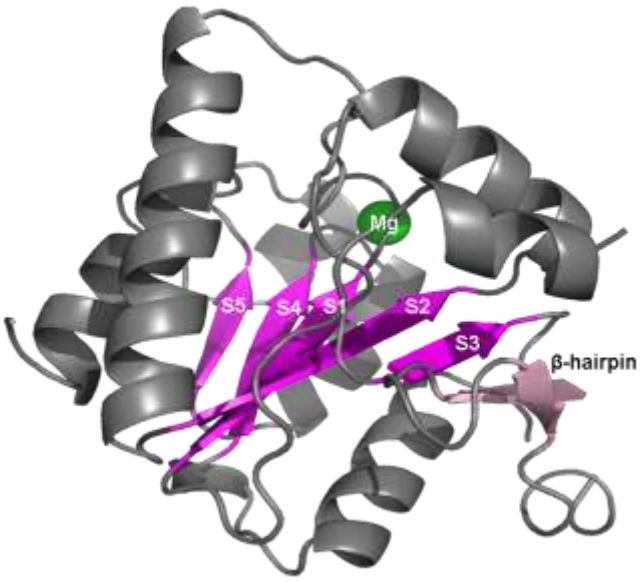


Figure 1.10: Identification of the structural characteristic of the core domain of NtdB, showing a Rossmann-like fold.
S1 to S5 are identified in magenta, β-hairpin in pink, and magnesium in green (PDB 3GYG).

Table 1.2: Amino acids corresponding to structural characteristics of the NtdB core domain

Structural characteristic	Amino acids
Rossmann-like fold	
S1	Y20 – D25
S2	L59 – T65
S3	F87 – S90
S4	A227 – G231
S5	N245 – L248
β-hairpin	T94-Y98 and G105-Q107

Apart from the Rossmann-like fold, some additional structural elements in the core can be observed through members of different subcategories of the HAD superfamily. One of these elements can be observed in NtdB (**Figure 1.10**) and consists

of a β-hairpin, found after S3, and in the same plane as the Rossmann-like fold. This structural feature is found in sucrose-phosphate phosphatases (SPP),[51] phosphomannomutases,[52] TPP,[53] mannosyl-3-phosphoglycerate,[54] and Cof-type phosphatases.[44]

1.5.3 Conserved amino acid residues in the core domain

Almost all members of the HAD superfamily contain four highly conserved motifs in the core domain and NtdB is no exception (**Table 1.3**).[44, 55]

Table 1.3: NtdB contains the four conserved motifs found in the core domain of HAD superfamily members

Conversed motif	Amino acids in NtdB
Motif 1 (DxD)	D25 – X – D27
Motif 2 (S/T)	T65
Motif 3 (K)	K209
Motif 4 (GDxxxD)	G231 – D232 – X – X – X – D236

Motif 1, found at the end of S1, consists of the amino acid sequence DxD (where x represents any amino acid). The first aspartate acts as a nucleophile and forms the aspartate-bound intermediate. This intermediate was indirectly proven with different members of the HAD superfamily, using radioactive (^{32}P) phosphate[56,57] and mass spectrometry,[58, 59] and directly observed in the crystal structure of β-phosphoglucomutase from *Lactococcus lactis*.[60] This first aspartate also helps to coordinate the binding of magnesium.[61-63] The second aspartate residue acts as a general acid-base for enzymes such as phosphatases, phosphotransferases,[53] and phosphomutases.[44] It can be replaced by alanine, for phosphonatases[61] or threonine, in the case of P-type ATPases.[62] Regardless of the amino acid, its carbonyl group also helps to coordinate Mg^{2+}.[61-63]

Motif 2 is highly conserved and contains either a serine or a threonine residue located at the end of S2. In the case of NtdB, this residue is a threonine. The side-chain

hydroxyl group is within hydrogen bonding distance of the phosphate oxygen, helping to orient it for nucleophilic attack by the aspartate.[64] This serine or threonine residue is also present to stabilize the intermediate formed.[44, 60]

Motif 3 is composed of a conserved lysine residue, which is located at the middle or the end (in the case of NtdB) of the helix after S4. Its function is to stabilize the negative charges of the intermediate.[44, 64] From crystal structures, such as sugar phosphatases[65], and phosphoserine phosphatase[62], it was observed that this residue forms a hydrogen bond with the first aspartate of Motif 1.[60, 61]

Motif 4 is a sequence containing two aspartates residues, either DD, GDxxxD, or GDxxxxD, located at the end of S4. These aspartates coordinate the Mg^{2+} in the active site,[38, 44, 61]and in some cases, via a water molecule.[53, 61] For NtdB, the motif is GDxxxD.

1.5.4 Secondary structure of the cap domain

The cap domain of members of the HAD superfamily can be divided into three categories: C0, C1 and C2.[44] NtdB possesses a C2 cap, inserted after S3 and characterized by a β-sheet, made up of at least three antiparallel β-strands. C2 caps are further subdivided into the HisB family, NagD family, and Cof hydrolase assemblage (which contains TPP and SPP, for example).[44]

Due to its structure, NtdB is believed to belong to the Cof hydrolase assemblage, its cap resembling other phosphate phosphatases of this group. NtdB contains a β-sheet composed of three antiparallel β-strands, as well as two α-helices in an αβαββ pattern (see **Table 1.4 and Figure 1.11**). This is similar to the secondary structure of a phosphoglycolate phosphatase from *Thermoplasma acidophilum* (1L6R), which is from the Cof-hydrolase assemblage.[64]

Table 1.4: Amino acids corresponding to structural characteristics of the NtdB cap domain

Cap motif	Amino acids
α_1	S120 – N136
β_1	F154 – Q157
α_2	E161 – Y178
β_2	S197 – I203
β_3	V180 – R186

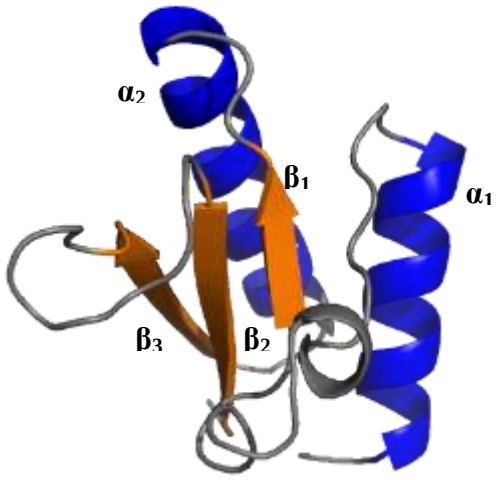


Figure 1.11: Identification of the structural characteristic of the cap domain of NtdB.
The β-strands are in orange and the α-helices in blue (PDB: 3GYG).

1.5.5 Conserved amino acid residues in the cap domain

The cap domain contains no specific amino acid motifs common to all members of the HAD superfamily. Instead, some commonalities can be found within specific sub-groups. For members possessing a C2 cap, a common specificity loop has been proposed.[66] This specificity loop serves to orient important residues, ensuring interaction with the leaving group of the substrate.[65] In the case of a phosphoglycolate phosphatase, the specificity loop positions three polar residues for interaction with the glycolate portion of the substrate.[64, 65] A more promiscuous enzyme, a sugar phosphatase from

Bacteroides thetaiotaomicron, has two specificity loops containing one polar and two non-polar residues, respectively.[65]

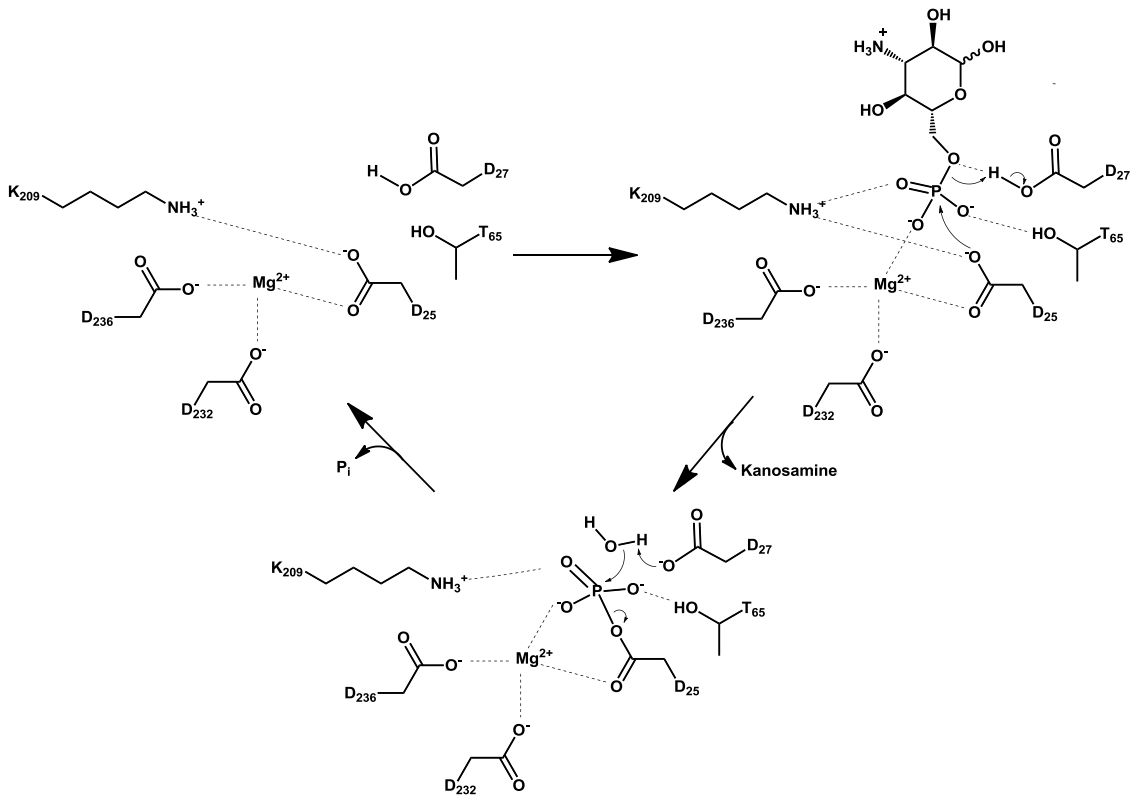
For NtdB, the crystal structure shows that this specificity loop contains 18 amino acids, R186 to I203. This is a relatively long loop compared to the other types of sugar phosphatases mentioned previously which contain only 9 or 10 residues.

1.6 Proposed role of NtdB

NtdB is proposed to act as a phosphatase and it possesses all the structural features of a phosphatase. In its core, not only does NtdB contain all four conserved motifs of members of the HAD superfamily, but it also contains features common to phosphatases such as the second aspartate of motif 1, and a β-hairpin, found after S3. NtdB possesses a C2 cap, common to members of the Cof family, and its αββαββ pattern is structurally closer to phosphatases than any other class of enzymes. Moreover, phosphatase activity has been observed with K6P.[67]

1.7 Proposed catalytic mechanism of NtdB

Based on the catalytic mechanism of members of the HAD superfamily and more specifically phosphatases, the following catalytic mechanism is proposed for NtdB (**Scheme 1.6**):



Scheme 1.6: Proposed catalytic mechanism of NtdB.

First, Mg^{2+} binds to the side chain carboxyl group of D25, D232, and D236 of the enzyme, as well as the backbone carbonyl of D27 (not shown), with or without the help of water molecules. The enzyme is in the open conformation, allowing the phosphorylated substrate, K6P, to access the active site located in the core. T65 helps to orient the phosphate group in the right position for the next step. The enzyme adopts a closed-conformation, excluding bulk solvent from the active site. D25, which forms a hydrogen bond with K209 in the open conformation, is in the right position to perform a nucleophilic attack on the phosphate group, consequently forming a phospho-enzyme intermediate.

The enzyme reverts back to its open conformation, liberating the dephosphorylated product, kanosamine, and allowing solvent to access the active site. The phospho-enzyme intermediate is stabilized by T65 and K209 until a water molecule,

deprotonated by D27, acts as a nucleophile on the phospho-enzyme intermediate. Inorganic phosphate is then liberated and the enzyme is returned to its original state.

2 RESEARCH PURPOSE

2.1 Research Hypothesis

The main hypothesis of this project was that NtdB belongs to the HAD superfamily and kanosamine-6-phosphate (K6P) is the native substrate. The active site residues would include D25, acting as a nucleophile, D27, performing the role of a general acid/base, T65 and K209, stabilizing the intermediate, and D232 and D236 coordinating the magnesium with D25. The cap domain would contain residues necessary for substrate binding.

A second hypothesis was that, based on the similarity of the sequence, KabB performs the same role as NtdB.

2.2 Research Objectives

The long-term goal of this project is to understand how 3,3'-neotrehalosadiamine (NTD) is produced. This thesis focuses on the enzyme NtdB of the *ntd* operon. The first research objective was the purification of NtdB. Once pure, the phosphatase activity of NtdB could be investigated and its substrate specificity examined. A second objective of this thesis was to perform site-directed mutagenesis in order to investigate the proposed catalytic mechanism of NtdB as well as identify the essential amino acids in the cap domain of NtdB. An important aspect of this thesis was to obtain kinetic constants for the reactions catalyzed by NtdB and its mutants.

As part of this research, KabB, an enzyme resembling NtdB was found. As part of this thesis, research on KabB was initiated. The research objectives related to KabB consisted of sub-cloning *kabB* in pet28b and expressing it in *E. coli*. The purification of KabB would allow to compared kinetics and substrate specificity between KabB and NtdB.

3 EXPERIMENTAL

3.1 Materials and instrumentation

3.1.1 Reagents and materials

Chemicals and reagents were purchased from Alfa Aesar (Ward Hill, MA), Bio-Rad (Mississauga, ON), EMD Millipore (Gibbstown, NJ), Sigma-Aldrich, Canada, Ltd. (Oakville, ON), TCI America (Portland, OR), or VWR (Mississauga, ON) and were of reagent grade or higher. DNA primers were synthesized by Alpha DNA (Montréal, QC). Restriction enzymes and their associated buffers were purchased from New England Biolabs (Pickering, ON). QIAprep® Spin Miniprep kits were purchased from QIAGEN, Inc. (Toronto, ON), E.Z.N.A.™ Plasmid Mini kits from Omega Bio-Tek Inc. (Norcross, GA), and EnzChek® Phosphate Assay kits from Invitrogen Canada, Inc. (Burlington, ON). For kinetics, NTD-6-phosphate, NTD-6'-phosphate, NTD-6,6'-bisphosphate and kanosamine-1-phosphate (K1P) were synthesized in Dr. Palmer's lab by Dr. Shazia Anjum and Dr. Rajendra Jagdhane (Department of Chemistry, University of Saskatchewan, Saskatoon, SK, Canada).

3.1.2 Plasmids, bacterial strains and growth conditions

The pUC18-*ntdABC* plasmid encoding the *ntdA*, *ntdB*, and *ntdC* genes was provided by Dr. Kozo Ochi (Microbial Function Laboratory, National Food Research Institute, Tsukuba, Ibaraki, Japan). The pET-28b-*ntdB* plasmid was previously prepared by David Langill (Department of Chemistry, University of Saskatchewan, Saskatoon, SK, Canada).[15] The pET-22b-*kabABC* and pET-15b-*kabB* plasmids were provided by Dr. Michael Thomas (Department of Bacteriology, University of Wisconsin-Madison, Madison, WI). Competent cells were generated from *E. coli* XL1-Blue and BL21-Gold® cell lines, purchased from Stratagene (la Jolla, CA). *E. coli* strains containing the desired

plasmid were routinely grown overnight at 37 °C in Luria-Bertani (LB) medium (with shaking at 250 rpm) or on LB-agar plates, either containing 50 µg/mL kanamycin or ampicillin. Plates were stored at 4 °C.

3.1.3 Instrumentation

Deionized water was purified using a Barnstead NANOpure® DIamond™ (UV/UF) Ultrapure water system to 18.2 MΩ. Liquids and materials were sterilized using either an Accusterilizer AS12 (VWR) or Castle® 233LS Vacuum steam sterilizer (Getinge). A Virsonic 600 ultrasonic cell disrupter was used to lyse cells. The purity of proteins was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on a Bio-Rad mini-PROTEAN® 3 electrophoresis system, and concentrations were determined using a NanoDrop® ND-100 spectrophotometer. An Eppendorf Mastercycler® Gradient thermal cycler was used for polymerase chain reaction (PCR). DNA separation and analysis was achieved using a Bio-Rad Mini-Sub® Cell GT agarose gel electrophoresis system. A Beckman Coulter® Avanti® J-E centrifuge and Beckman Microfuge® 18 were used for DNA purification, cell harvesting, and protein sample preparation. A PiStar-180 spectrometer (Applied Photophysics) was used for circular dichroism (CD) experiments, where the temperature was controlled by an Isotemp 3016D (Fisher Scientific) circulating water bath. Spectrophotometric assays were accomplished using a Beckman DU-640 UV/Vis spectrophotometer, connected to a circulating water bath to control the temperature. DNA sequencing was performed by the DNA Technologies Unit of the National Research Council, Plant Biotechnology Institute (Saskatoon, SK, Canada). The Saskatchewan Structural Sciences Center (Saskatoon, SK, Canada) carried out mass spectrometric analysis of NtdB using an API Qstar® XL MS/MS system with electrospray ionization source (ESI-Q-TOF) (Applied Biosystems, Toronto, ON). Nuclear magnetic resonance was performed with a Bruker Avance 500 MHz UltraShield™ NMR spectrometer. Spectra were calibrated at 4.79 ppm with residual water in D_2O.

Protein sequences were found using GenBank®[68] or Basic Local Alignment Search Tool (BLAST).[45] Protein sequence alignments were generated with *ClustalW*[20], using the Blosum scoring matrix[21]: opening gap penality = 10; extending gap penality = 0.20; distance gap penality = 5. The sequence alignment images were created by *ESPript* V2.2[22]. Protein properties were calculated using the software Protein Calculator 3.3 (http://www.scripps.edu/~cdputnam/protcalc.html) from The Scripps Research Institute (La Jolla, CA). The software Oligonucleotide Properties Calculator (OligoCalc)[69] was used to generate primer sequences and NEBcutter v2.0[70] to introduce restriction sites. Using chain A of NtdB (PDB 3GYG) as a template, homology modeling of KabB was performed using SWISS-MODEL.[71,72,73]

3.2 Chemical synthesis

3.2.1 Synthesis of K6P and Glucosamine-6-phosphate

The synthesis of K6P and glucosamine-6-phosphate (GlcAm6P) followed the procedure of Guo and Frost.[24] Kanosamine or glucosamine (110 mg, 0.61 mmol), ATP (500 mg, 1.00 mmol), citric acid (100 mg, 0.48 mmol), and $MgCl_2$ (250 mg, 1.23 mmol) were dissolved in 35 mL distilled water, adjusted to pH 8.0 by the addition of aqueous NaOH, and degassed for 10 minutes with nitrogen. To the solution, 6.3 mg (248 U) of hexokinase (from baker's yeast) was added and stirred at room temperature. The pH was monitored periodically and maintained between 7.5 and 8.0 by the addition of 1 M NaOH. After 24 hours of stirring, the hexosephosphate was purified from the reaction mixture by anion-exchange chromatography on AG® 1-X8 resin (acetate form, Bio-Rad) and eluted by sequentially adding 50 mL of increasing concentrations of acetic acid (0 N, 0.5 N, 1.0 N, 1.5N, 2.0 N). Ten mL fractions were collected and freeze-dried.

Fractions containing the purified hexosephosphate were determined by ^{1}H NMR and compared to literature values.[24,74] **K6P**: ^{1}H NMR: δ ppm (500 MHz, D_2O) 5.12 (d, 1H), 4.57 (d, 1H), 3.95 (m, 2H), 3.84 (m, 1H), 3.78 (m, 2H), 3.63 (m, 2H), 3.44 (d, 1H),

3.25 (m, 2H), 3.01 (t, 1H). ^{31}P NMR: δ ppm (200 MHz, D_2O) 1.80, 1.66. **GlcAm6P**: ^{1}H NMR: δ ppm (500 MHz, D_2O) 5.38 (d, 1H), 4.85 (d, 1H), 4.00 (m, 2H), 3.90 (m, 1H), 3.84 (t, 1H), 3.59 (m, 2H), 3.51 (m, 2H), 3.25 (dd, 2H), 2.94 (dd, 1H).

3.3 Biochemical procedures

3.3.1 Buffers

The buffers used in this thesis are listed alphabetically in **Table 3.1**. Competent cells were prepared with sterile Transformation buffer. For the protein purification protocol, filtered buffers B, W, and E were used, corresponding to the binding, wash and elution steps, respectively. Dialysis was performed either in dialysis buffer (for storage and kinetics) or CD buffer (CD experiments). All the kinetic assays were done in Universal buffer.

Table 3.1: Content and pH of various buffers

Buffers	pH	Content
B	7.9	20 mM HEPES-HCl, 500 mM NaCl, 5mM imidazole
CD	7.8	10 mM potassium phosphate, 15% glycerol
Dialysis	7.5	20 mM HEPES-HCl, 500 mM NaCl
E	7.9	20 mM HEPES-HCl, 500 mM NaCl, 55 mM imidazole
SDS	8.7	25 mM Tris, 250 mM Glycine, 0.1% SDS
TAE	8.0	40 mM Tris-acetate, 1mM EDTA
Transformation	7.5	25 mM Tris-HCl, 57 mM $CaCl_2$
Universal	Various	100 mM acetic acid, 50 mM Bis-Tris, 50 mM TEA
W	7.9	20 mM HEPES-HCl, 500 mM NaCl, 30 mM imidazole

Abbreviations: Bis-Tris, bis(2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane; EDTA, ethylenediaminetetraacetic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; SDS, Sodium dodecyl sulfate; TAE, tris-Acetate-EDTA; TEA, triethanolamine;

3.3.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

SDS-PAGE was performed using acrylamide-bisacrylamide gels and a Tris-glycine buffer (SDS buffer) system as described by Sambrook and Russel.[75] The resolving gel and the stacking gel had acrylamide-bisacrylamide concentrations of 12% and 4.5% respectively. Gels were electrophoresed at 90 V for 15 min followed by 200 V for 30 min at room temperature.

To visualize the proteins, gels were stained by soaking in Fairbanks A (0.05% w/v coomassie brilliant blue, 25% v/v isopropanol, 10% v/v glacial acetic acid), heating for 30 seconds in a microwave (700 watts) followed by gentle rocking at room temperature for 5 min. Gels were then rinsed with distilled water. To destain the gels, the same procedure was applied, using successively Fairbanks B (0.005% w/v coomassie brilliant blue, 10% v/v isopropanol, 10% v/v glacial acetic acid), Fairbanks C (0.002% w/v coomassie brilliant blue, 10% v/v glacial acetic acid), and Fairbanks D (10% v/v glacial acetic acid).

3.3.3 Preparation of BL21-Gold® and XL1-Blue competent cells

E. coli BL21-Gold® competent cells, and *E. coli* XL1-Blue competent cell, were prepared as follows: 5 mL of LB broth containing 50 µg/mL of tetracycline was inoculated with frozen glycerol stock of the desired strain and incubated at 37 °C and 250 rpm overnight. The overnight culture (1 mL) was used to inoculate 150 mL of LB broth with tetracycline (30 µg/mL) and incubated at 37 °C with shaking at 250 rpm until achieving an optical density at 600 nm of approximately 0.4. The cell culture was harvested by centrifugation at 1 500 *x g* for 20 minutes at 4 °C. The cell pellet was resuspended in 20 mL ice-cold transformation buffer and incubated on ice for 30 minutes. The resuspended cells were pelleted as before and resuspended in 5 mL ice-cold transformation buffer. The competent cells were diluted with sterile, ice-cold 80%

glycerol to a final concentration of 15% glycerol and aliquots of 100 µL were stored at -80 °C until used.

3.3.4 Transformation

A 100 µL frozen aliquot of BL21-Gold® or XL1-Blue competent cell was thawed on ice before adding 2 µL of the desired plasmid. The transformation mixture was incubated on ice for 30 minutes, followed by 42 °C for one minute, and placed back on ice for two minutes. The transformation mixture was used to inoculate 900 µL of pre-warmed LB broth and incubated at 37 °C with shaking at 250 rpm for one hour. One hundred µL of the culture was spread onto a pre-warmed LB-agar plate containing 50 µg/mL kanamycin. The plate was allowed to incubate overnight at 37 °C, yielding several colonies.

3.3.5 Plasmid isolation

E. coli XL-1 Blue cells containing the desired plasmid were grown overnight in 5 mL LB broth containing 50 µg/mL kanamycin, at 37 °C with agitation at 250 rpm. The cell culture was harvested by centrifugation at 18 000 *x g* for 3 minutes at ambient temperature. Plasmid was isolated in distilled water using a QIAprep® Spin Miniprep kit or E.Z.N.A.™ Plasmid Mini kit, according to the manufacturer's protocol.

3.3.6 Polymerase chain reaction

Polymerase chain reaction (PCR) was performed using a KAPA HiFi™ PCR kit. Each reaction contained 50 ng of desired plasmid, 0.3 mM deoxyribonucleotide triphosphate mix (equal amounts of each dATP, dCTP, dGTP, and dTTP), 0.8 µM of forward and 0.8 µM reverse primers of the desired mutation, 1X KAPA HiFi™ buffer, and 1 U KAPA HiFi™ DNA polymerase (25 µL final volume). A negative control without DNA polymerase was also performed. The PCR cycling protocols all started with

an initial denaturation at 95 °C for 2 or 5 minutes which was followed by cycling three steps: denaturation, annealing and polymerization. The protocol ended with a final extension at 72 °C for 3 minutes.

To remove the parental DNA, the amplification products were treated with *Dpn*I and incubated at 37 °C for 1 hour. A positive control without *Dpn*I was also performed. *E. coli* XL1-Blue competent cells and *E. coli* BL21-Gold® competent cells were transformed with the resulting mixture using the transformation protocol previously described. Colonies from the plates were grown overnight in 5 mL LB broth containing 50 µg/mL kanamycin, at 37 °C with agitation at 250 rpm. The plasmids were isolated as described in the previous subsection and the presence of the desired sequence in the plasmids was confirmed by nucleotide sequencing.

3.4 Molecular biology

3.4.1 Site-directed mutagenesis of the *ntdB* gene

NtdB was previously sub-cloned into pET-28b.[15] The plasmid encodes a hexahistidine tag at the N-terminus of the protein. Different mutations in the cap domain (R149K, H152F, C187A, D193N, D196N, D199N, D201N) and the core domain (D25N, D27A, D27N, T65A, T65S, K209E, K209R, D232A, D236A) of NtdB were introduced **(Table 3.2)**. (Note: unless otherwise specified, mutants will be referred to by their amino acid substitution, and refer to NtdB.) Restriction sites were introduced in some mutants with the addition of a silent mutation. D27A and D27N contained a *Pvu*I restriction site, T65A contained a *Sac*I restriction site, and K209E and K209R contained an *Acc*I restriction site. No convenient silent mutations leading to the addition of a useful restriction site could be created for the remaining mutations. D25N contained three silent mutations in order to increase the GC content of the primers. The pET-28b-*ntdB* plasmid (isolated as described above) was used as a template for the mutagenesis reaction and the mutagenic primers are listed in **Table 3.3**.

Table 3.2: *ntdB* mutations

Mutation	Base change	Restriction site
D25N	69 T → C*	None
	73 G → A	
	78 T → C*	
	93 T → C*	
D27A	80 A → C	102 A → G *Pvu*I
D27N	79 G → A	102 A → G *Pvu*I
T65A	193 A → G	177 A → C *Sac*I
T65S	194 C → G	None
R149K	445 C → A	None
	446 G → A	
H152F	454 C → T	None
	455 A → T	
C187A	559 T → C	None
	560 G → C	
D193N	577 G → A	None
D196N	586 G→ A	None
D199N	595 G → A	None
D201N	601 G → A	None
K209E	626 A → G	603 T → C *Acc*I
K209R	625 A → G	603 T → C *Acc*I
D232A	695 A → C	None
D236A	707 A → C	None

Table 3.3: Mutagenesis primers of *ntdB* and their melting temperature (T_m)

Mutant	Primer	Oligonucleotide sequence (5' → 3')	T_m (°C)
D25N	Forward	GTTTTcTGTaATTTcGATGAAACGTATTTcCCGC	60.8
	Reverse	GCGGgAAATACGTTTCATCgAAATtACAgAAAAC	
D27A	Forward	GTGATTTTGcTGAAACGTATTTTCCGCATACgATCGAC	64.5
	Reverse	GTCGATcGTATGCGGAAAATACGTTTCAgCAAAATCAC	
D27N	Forward	GTGATTTTaATGAAACGTATTTTCCGCATACgATCGAC	62.2
	Reverse	GTCGATcGTATGGGAAAATACGTTTCATtAAAATCAC	
T65A	Forward	GAGCTcATAATCGGATGGGTTgCTGGGAG	64.3
	Reverse	CTCCCAGcAACCCATCCGATTATgAGCTC	
T65S	Forward	GCTAATAATCGGATGGGTTAgTGGGAGCAGTATAG	64.4
	Reverse	CTATACTGCTCCCAcTAACCCATCCGATTATTAGC	
R149K	Forward	CCTCAAACCCAATTAGGAAAGTCAaaGTATAAGC	60.8
	Reverse	GCTTATACttTGACTTTCCTAATTGGGTTTGAGG	
H152F	Forward	GGAAAGTCACGGTATAAGttTAACTTCTACTATCAGGAAC	63.5
	Reverse	GTTCCTGATAGTAGAAGTTAaaCTTATACCGTGACTTTCC	
C187A	Forward	CAGTAAATATAAATCGGccTAATCCCTTAGCAGG	62.0
	Reverse	CCTGCTAAGGGATTAggCCGATTTATATTTACTG	
D193N	Forward	GTAATCCCTTAGCAGGCaATCCAGAAGACAG	63.0
	Reverse	CTGTCTTCTGGATtGCCTGCTAAGGGATTAC	
D196N	Forward	GCAGGCGATCCAGAAaACAGCTATGATGTAG	63.0
	Reverse	CTACATCATAGCTGTtTTCTGGATCGCCTGC	
D199N	Forward	GCAGGCGATCCAGAAGACAGCTATaATGTAG	63.0
	Reverse	CTACATtATAGCTGTCTTCTGGATCGCCTGC	
D201N	Forward	GACAGCTATGATGTAaATTTTATCCCCATAGGGACAGG	64.5
	Reverse	CCTGTCCCTATGGGGATAAAATtTACATCATAGCTGTC	
K209E	Forward	GTAGAcTTTATCCCCATAGGGACAGGAgAGAATG	64.4
	Reverse	CATTCTcTCCTGTCCCTATGGGGATAAAgTCTAC	
K209R	Forward	GTAGAcTTTATCCCCATAGGGACAGGAAgGAATG	64.4
	Reverse	CATTCcTTCCTGTCCCTATGGGGATAAAgTCTAC	
D232A	Forward	GCTATCGCATTTGGGGcTAGTGGAAATGATGTTC	64.4
	Reverse	GAACATCATTTCCACTAgCCCCAAATGCGATAGC	
D236A	Forward	GTGGAAATGcTGTTCGTATGTTACAGACAG	60.3
	Reverse	CTGTCTGTAACATACGAACAgCATTTCCAC	

Unchanged nucleotides are upper case, while mismatched nucleotides are lower case, and restriction sites are underlined.

Table 3.4: PCR cycling protocol for the mutagenesis of *ntdB*

Mutation	Initial denaturation	Denaturation	Annealing	Polymerization and extension	Cycles
D25N	95 °C 300 s	98 °C 20 s	56 °C to 64 °C 15 s	72 °C 180 s	30
D27A	95 °C 120 s	98 °C 20 s	55 °C to 65 °C 30 s	72 °C 180 s	16
D27N	95 °C 120 s	98 °C 20 s	55 °C to 65 °C 30 s	72 °C 180 s	16
T65A	95 °C 120 s	98 °C 20 s	55 °C to 65 °C 30 s	72 °C 180 s	16
T65S	95 °C 300 s	98 °C 30 s	55 °C to 65 °C 60 s	72 °C 180 s	16
R149K	95 °C 300 s	98 °C 20 s	58 °C to 63 °C 15 s	72 °C 180 s	25
H152F	95 °C 300 s	98 °C 20 s	60 °C to 65 °C 15 s	72 °C 180 s	30
C187A	95 °C 300 s	98 °C 20 s	58 °C to 63 °C 15 s	72 °C 180 s	25
D193N	95 °C 300 s	98 °C 20 s	60 °C to 65 °C 15 s	72 °C 180 s	30
D196N	95 °C 300 s	98 °C 20 s	60 °C to 64 °C 60 s	72 °C 180 s	16
D199N	95 °C 300 s	98 °C 20 s	60 °C to 65 °C 15 s	72 °C 180 s	30
D201N	95 °C 300 s	98 °C 30 s	55 °C to 65 °C 60 s	72 °C 180 s	16
K209E	95 °C 300 s	98 °C 30 s	55 °C to 65 °C 60 s	72 °C 180 s	16
K209R	95 °C 300 s	98 °C 30 s	55 °C to 65 °C 60 s	72 °C 180 s	16
D232A	95 °C 300 s	98 °C 20 s	56 °C to 64 °C 15 s	72 °C 180 s	30
D236A	95 °C 300 s	98 °C 30 s	56 °C to 60 °C 60 s	72 °C 180 s	16

The PCR protocol followed the methodology described before, using four different cycling protocols (**Table 3.4**).

3.4.2 KabB

3.4.2.1 Polymerase chain reaction amplification

One colony of pET-22-*kabABC* was used to inoculate 5 mL LB broth containing 50 µg/mL ampicillin which was grown overnight at 37 °C, with agitation at 250 rpm. The cell culture was pelleted by centrifugation at 18 000 *x g* for 3 minutes at room temperature, and the plasmid was isolated using an E.Z.N.A.™ Plasmid Mini kit. Amplification of the *kabB* gene was performed using PCR, following the same procedure outlined previously (§3.3.6), using pET-22-*kabABC*. The initial denaturation was performed at 95 °C, for 300 seconds, followed by 28 cycles consisting of: denaturation (98 °C, 20 seconds), annealing (55-57 °C, 15 seconds), and extension (72 °C, 90 seconds). The *kabB* primers used for PCR amplification can be found in **Table 3.5**. The primers introduced two restriction sites: *Nde*I in the 5′ end and *Bam*HI in the 3' end of the

gene. The PCR product was purified from the reaction mixture following the QIAquick® PCR purification kit protocol from Qiagen.

Table 3.5: Primers for *kabB* PCR amplification

	Primer	Oligonucleotide sequence (5′ → 3′)	T_m (°C)
KabB	Forward	GGAGGGATTAAcatATGTCAATAAGTAATGAAG	58.2
	Reverse	CATTATATGAggATCCCCTTTAATTAATAATGTG	57.3

Unchanged nucleotides are upper case, while mismatched nucleotides are lower case, and restriction sites are underlined.

3.4.2.2 Restriction digest of pET-28b and *kabB*

One microgram of pET-28b plasmid or *kabB* PCR amplified fragments was incubated in NEBuffer #3 (5 µL) containing Bovine Serum Albumin (10 µg) and *Nde*I (30 U) for one hour, at 37 °C. *Bam*HI (20 U) was added to the mixture, which was then incubated for an additional hour (final volume 50 µL). Restriction enzymes and fragments less than 100 bp were removed using a QIAquick® PCR purification kit.

3.4.2.3 Ligation of *kabB* into pET-28b

The ligation reaction was prepared at a 3:1 insert-to-vector molar ratio. The ligation mixture contained 15 fmol of *kabB* PCR amplified fragments, 5 fmol pET-28b plasmid, and T4 DNA ligase (20 U) in 1X T4 DNA ligase buffer (final volume 10 µL). The ligation mixture was incubated overnight at 16 °C. The ligation mixtures were used to transform XL1-Blue competent cells as described above. Plasmids were isolated from the colonies of transformed cells using the methodology mentioned previously. To confirm successful ligation, a restriction digest using *Sap*I was performed and the products were visualized on a 1% agarose gel containing GelRed™ dye (Biotium Inc.) that had been developed in TAE buffer at 120 V. The presence of the intact *kabB* gene in plasmids containing a *Sap*I restriction site was confirmed by nucleotide sequencing.

3.4.3 KabB mutant

One mutation in the cap domain of KabB (N201D) was performed, which consisted of a single base pair change at position 601, from adenine to guanine. The pET-28b-*kabB* plasmid (isolated as described above) was used as a template for the mutagenesis reaction and the mutagenic primers are listed in **Table 3.6**.

Table 3.6: Mutagenesis primers of KabB N201D and their melting temperatures

	Primer	Oligonucleotide sequence (5' → 3')	T_m (°C)
N201D	Forward	GCTGGTGATCCAGAAgATTGCTATGATGTAG	61.7
	Reverse	CTACATCATAGCAATcTTCTGGATCACCAGC	

Unchanged nucleotides are upper case, while mismatched nucleotides are lower case, and restriction sites are underlined.

The PCR protocol followed the methodology described before. The initial denaturation was performed at 95 °C, for 300 seconds, followed by 16 cycles consisting of: denaturation (98 °C, 20 seconds), annealing (60-64 °C, 60 seconds), and extension (72 °C, 180 seconds). DNA sequencing confirmed the presence of the desired mutation.

3.5 Protein expression and purification

3.5.1 Growth and induction of the desired plasmid

One colony of the desired plasmid was used to inoculate 5 mL LB broth, containing 50 µg/mL kanamycin, and incubated at 37 °C, with agitation at 250 rpm overnight. Of the overnight culture, 2 mL were used to inoculate 500 mL of LB broth containing kanamycin (50 µg/mL), which was then incubated at 37 °C, shaking at 250 rpm, until an optical density at 600 nm of approximately 0.5 was obtained. The cells were induced with the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and incubated for 2 hours at 37 °C, with agitation at 250 rpm. The cell culture was harvested

by centrifugation at 5 000 *x g* for 30 minutes at 4 °C and two aliquots of approximately 0.4 g of wet cells were frozen at -20 °C until used.

3.5.2 Purification of NtdB, KabB, and mutants

One aliquot of frozen cells (§3.5.1) was resuspended in 5 mL buffer B and lysed by sonication at power 3, carrying out 36 cycles of 5 seconds pulse and 20 seconds rest. The cell lysate was centrifuged at 20 000 *x g* and 4 °C for 20 minutes. Since all the proteins contained a His-Tag, the supernatant was filtered using a 0.45 µM syringe filter and purified by a Ni^{2+}-affinity chromatography with a 1 mL HiTrap™ IMAC FF column (GE Healthcare), which was pre-treated with 100 mM $NiSO_4$ and equilibrated with buffer B. The sample was loaded and the column was washed with buffer B (10 column volumes) and buffer W (10 column volumes). The protein was then eluted with buffer E (10 column volumes). Column fractions were analyzed by SDS-PAGE, and those containing the desired enzyme were combined and dialyzed twice in a period of 24 hours against 1 L dialysis buffer. The protein was stored at -80 °C in 25% glycerol.

3.6 Circular dichroism

Pure wild type (WT) NtdB, KabB, and mutants were dialyzed against 1 L CD buffer, twice in a period of 24 hours. Far-UV CD measurements were performed at a temperature of 4 °C. The CD instrument was calibrated with a solution of 4 mM (1*S*)-(+)-10-camphorsulfonic acid at 290.5 nm. CD spectra were recorded from 190-260 nm in a 0.5 mm cell, using a 6 nm bandwidth, and a scan speed of 5 nm min^{-1}. Each spectrum was generated from the average of 4 scans. CD spectra were corrected with respect to the baseline, corresponding to the CD buffer, and subjected to curve-smoothing. The molar ellipticity ([θ]) was calculated from the ellipticity given by the instrument (θ), the molecular weight of the protein (MW), the protein concentration (C), the cell pathlength (l) and the number of amino acids in the protein (N_A), according to equation **3.1**[76]. CDNN deconvolution software[77] version 2.1 was used to estimate secondary structure.

$$[\theta]=\frac{100\times\theta\times MW}{C\times l\times N_A} \quad \textbf{(3.1)}$$

3.7 Kinetics

Kanosamine-6-phosphate (K6P) stock solutions were prepared weekly and contained sodium acetate as an impurity (confirmed by HPLC).[17]

3.7.1 *para*-Nitrophenylphosphate assay

The assay mixture contained universal buffer and 20 mM *para*-nitrophenylphosphate. After 10 minutes incubation at ambient temperature, 500 nM NtdB, KabB, or mutants were added to the solution for a final volume of 1 mL. The formation of *para*-nitrophenol was monitored by spectrophotometer at 410 nm, and its concentration calculated using Beer's law (**3.2**) The extinction coefficient of *para*-nitrophenylphosphate is 18 400 M^{-1} cm^{-1}.[78]

3.7.2 Malachite green assay

To prepare the Malachite Green solution, 250 mL of 4.2% (w/v) ammonium molybdate in 4 M HCl was added to 750 mL of 0.045% (w/v) Malachite Green for a final volume of 1 L.[79] The solution was stirred for 30 minutes and filtered through a 0.22 μm filter and stored at 4 °C for up to six months. Tween 20 (0.1%, v/v) was added to aliquots of filtered Malachite Green solution before use.

The assay mixture contained 25 nM NtdB, 10 mM $MgCl_2$, and universal buffer. The reaction was initiated with the addition of various concentrations of K6P for a final reaction volume of 250 μL. The reaction was stopped at different times with the addition of 500 μL of malachite green solution, and incubated at room temperature for 45 minutes.

Each assay was performed in triplicate. The reaction mixture without enzyme was used as a blank.

The formation of the phosphomolybdate-malachite green complex was monitored at 620 nm. Quantification of phosphate released from the reaction was accomplished using a standard curve generated using fixed concentrations of phosphate.

3.7.3 EnzChek® Phosphate assay

The kinetic assays followed the procedure of the EnzChek® Phosphate Assay kit, with modifications in the content of the assay mixture. The assay mixture contained universal buffer, 10 mM $MgCl_2$, 200 µM 2-Amino-6-mercapto-7-methylpurine ribonucleoside (MESG), one unit of purine-nucleoside phosphorylase (PNP), and various concentrations of substrate. The mixture was incubated at room temperature for 10 minutes. The coupled reaction was initiated by the addition of 25 nM, 250 nM, or 500 nM of enzyme (1 mL final volume). Each assay was duplicated.

The formation of 2-Amino-6-mercapto-7-methylpurine (AMMP) was monitored at 360 nm, which corresponded to the formation of inorganic phosphate from the reaction. The increase in concentration of P_i (ΔC) was calculated the Beer's law (**3.2**).

$$\Delta C = \frac{\Delta A}{\varepsilon \times l} \qquad \textbf{(3.2)}$$

where ΔA represents the change in absorbance per second approximated from the slope of the initial segment of the reaction curve, ε is the extinction coefficient of AMMP (11 000 M^{-1} cm^{-1})[80], and l is the path length (1 cm).

3.7.4 Data analysis

The kinetic data was fitted by non-linear regression to the Michaelis-Menten equation (**3.3**) using Sigma Plot v.11.0 software (Systat Software, Inc.),

$$\boldsymbol{V_o = V_{max}\ \frac{[S]}{K_m+[S]}} \qquad \textbf{(3.3)}$$

where V_0 is the initial velocity, which is the change in concentration per second (ΔC), V_{max} the maximum velocity, [S] the substrate concentration, and K_m the Michaelis-Menten constant for the substrate.

The k_{cat} value was calculated from V_{max} and $[E]_0$ according to the **Equation 3.4**

$$\boldsymbol{k_{cat} = \frac{V_{max}}{[E]_0}} \qquad \textbf{(3.4)}$$

where [E] corresponds to the total concentration of NtdB, KabB, or mutants in the reaction.

4 RESULTS AND DISCUSSION

4.1 Purification of NtdB

Subcloning of *ntdB* has previously been described.[15] The successful overexpression of *ntdB* in *E. coli* strain can be observed in **Figure 4.1**. For *in vitro* analysis of NtdB, the enzyme needed to be purified to homogeneity from the crude cellular extract. To avoid aggregation and obtain solubility, it was recommended to use a pH at least one or two units away from the isoelectric point of the protein. Using the software Protein Calculator 3.3, the isoelectric point of NtdB was calculated to be 6.16.

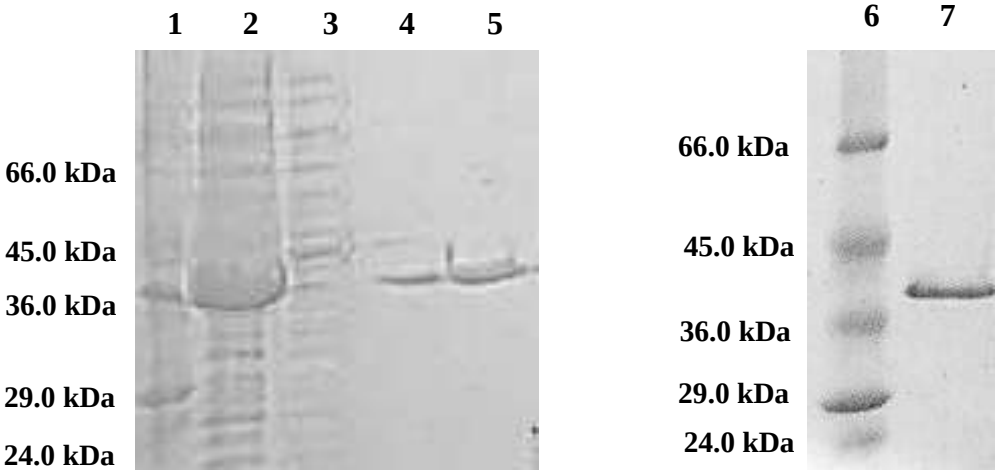


Figure 4.1: SDS-PAGE showing purification steps of WT NtdB.
Lane 1: Ladder. Lane 2: Crude WT NtdB. Lane 3: Flow-through. Lane 4: Wash. Lane 5: Eluted WT NtdB. Lane 6: Ladder. Lane 7: Pure WT NtdB.

NtdB contains a His-tag consisting of six histidine residues at the N-terminus. When the enzyme is passed through a chromatographic column containing Ni^{2+}, the His-tag chelates to the nickel, binding the enzyme to the column. All other non-tagged components flow through the column. We can observe in **Figure 4.1** that the flow-through contains very little, if any, NtdB. The wash step removed any loosely bound components and a small amount of NtdB can be perceived on the SDS-PAGE gel. When eluted, NtdB is the only component left.

The SDS-PAGE of pure WT NtdB was run next to a molecular weight ladder of known proteins (**Figure 4.1,** lane 6-7). Visually, the mass of NtdB appears to be slightly higher than 36 kDa. The expected mass of NtdB, based on its amino acid sequence including its His-tag, was 34 732 Da. This small discrepancy in the molecular weight of NtdB could be due to the accuracy of the gel which is within the range of 5-10%.[81,82] There are several examples of proteins of similar molecular weight to NtdB that also show this unexpected behavior by SDS-PAGE. For example, pepsin (35 500 Da) has been shown to appear higher in molecular weight, while carboxypeptidase A (34 400 Da) appears lower.[82]

4.2 Mass spectrometry of NtdB

To confirm the actual weight of NtdB, mass spectrometry was used. **Figure 4.2** shows two molecular weights for NtdB. The smaller peak, observed at 34 733 ± 2 Da, corresponds to the expected theoretical weight of NtdB. The second peak was observed at 34 602 ± 2 Da, a difference of 131 ± 3 Da, which is likely due to loss of the first methionine residue of the NtdB amino acid sequence.

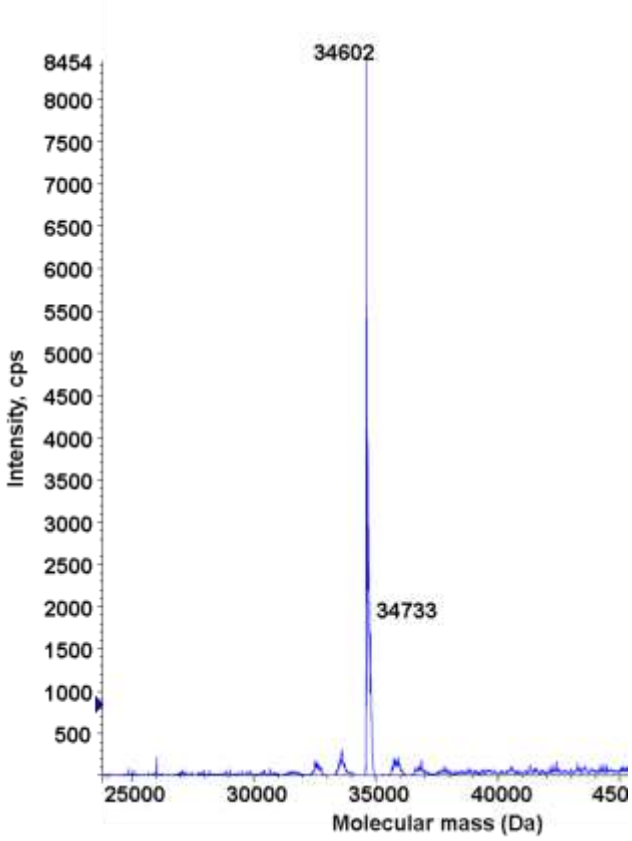


Figure 4.2: Molecular mass of NtdB obtained by mass spectrometry.

In *E. coli* cytosolic proteins, it is not uncommon for this first methionine residue to be excised, and the identity of the second residue seems to affect the likelihood of this phenomena. When comparing over hundred *E. coli* protein sequences with different second residues, a strong positive correlation was observed between the presence of glycine, alanine, serine, cysteine, proline, or threonine residues, and the removal of methionine.[83] Since an *E. coli* host was used to express His-tag NtdB and that the second residue is a glycine, it gives credence to the hypothesis that the mass of 34 602 Da observed in the mass spectrum comes from the excision of the first methionine residue.

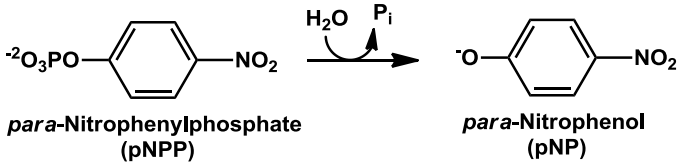
Nevertheless, mass spectrometric analysis confirms that the molecular mass of NtdB is correct, despite the inaccuracy of the SDS-PAGE gel.

4.3 Kinetic assays of NtdB

NtdB possesses all the physical characteristics of a HAD superfamily phosphatase (§1.5), such as the presence of a C2 cap domain, a core domain which contains a Rossmann-like fold, and several conserved amino acid residues, and motifs. Phosphatases remove a phosphate group from its substrate and the activity of such enzymes can be determined by different spectrophotometric methods, which measures consumption of substrate and/or production of product, or release of inorganic phosphate.

4.3.1 *para*-Nitrophenylphosphate assay

For more than 60 years, *para*-nitrophenylphosphate has been used as a substrate to easily assess the activity of bacterial phosphatases.[84] When *para*-nitrophenylphosphate is dephosphorylated, the yellow product, *para*-nitrophenol, can be spectrophotometrically monitored between 390 nm and 410 nm (**Scheme 4.1**).



Scheme 4.1: The hydrolysis of *para*-nitrophenylphosphate

This non-specific substrate has been previously used on phosphatases structurally similar to NtdB: a phosphatase from *Thermotoga maritima* (1NF2);[85] a phosphoglycolate phosphatase and trehalose-6-phosphate phosphatase from *Thermoplasma acidophilum* (1L6R and 1U02 respectively);[53, 64, 65] and a sugar phosphatase from *Bacteroides thetaiotaomicron* (1YMQ).[65] **Table 4.1** summarizes the kinetic values reported in the literature using *para*-nitrophenylphosphate as a substrate.

Table 4.1: Kinetic values of phosphatases structurally similar to NtdB using *para*-nitrophenylphosphate as a substrate.

PDB	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
1NF2[85]	0.10	16	6.3
1L6R[64]	0.58	3.1	190
1U02[53]	0.80	17	47
1YMQ[65]	0.083	0.77	110

Based on the results reported with other phosphatases, 20 mM of *para*-nitrophenylphosphate was used to test 500 nM NtdB. Activity was observed, confirming that NtdB act as a phosphatase (**Figure 4.3**.). The rate of the reaction was 0.0017 µM s^{-1} and this extremely low activity would indicate that the specificity of NtdB is different from other phosphatases, even those with a similar structure.

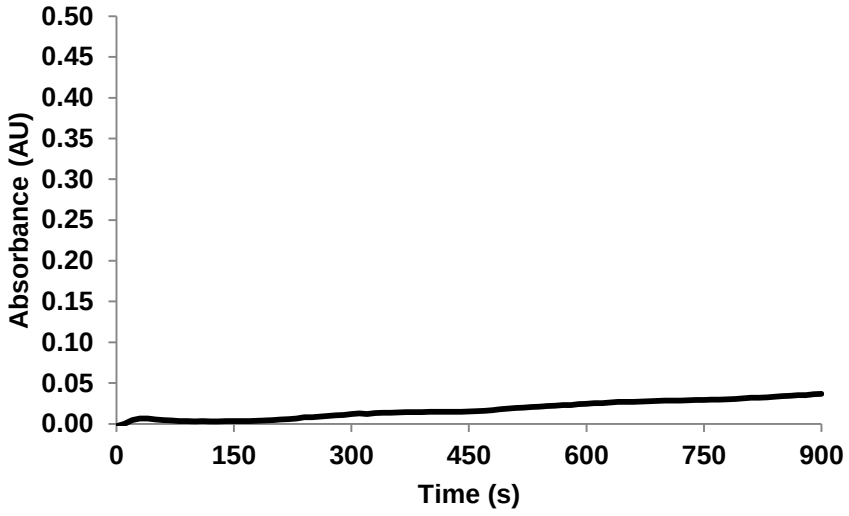


Figure 4.3: Hydrolysis of *para*-nitrophenylphosphate catalyzed by NtdB

It was concluded that the *para*-nitrophenylphosphate is a very poor substrate of NtdB, and therefore not a useful tool to determine the activity of the enzyme. This assay was not pursued.

4.3.2 Malachite green assay

Since not all dephosphorylated substrates can be easily observed using a spectrophotometer, another type of assay can be used, monitoring the release of inorganic phosphate from the phosphatase reaction. In 1946, Soyenkoff published a method of phosphate determination based on the formation of an insoluble dye-phosphomolybdate complex, causing a color change of the dye.[86] This method has been improved since, and presently, malachite green is the most commonly used dye, where the formation of the complex can be observed between 620 nm and 650 nm.[87] The malachite green assay has been used to establish the kinetic parameters of many phosphatases throughout the years.[64, 79, 88]

When tested with NtdB, the malachite green assay was successful, but limited (**Figure 4.4**). Using this assay, the results showed a very large variance. It is a discontinuous assay, whereby the reaction is stopped at precise time points (at least five) and the product, in this case phosphate, is measured. If this process is not automated, there is the possibility of a greater margin of error, which can be partially compensated for by repeating each experiment many times. These two factors make this assay a great consumer of substrates which were in limited quantity. This assay was abandoned in favor of the EnzChek® Phosphate assay.

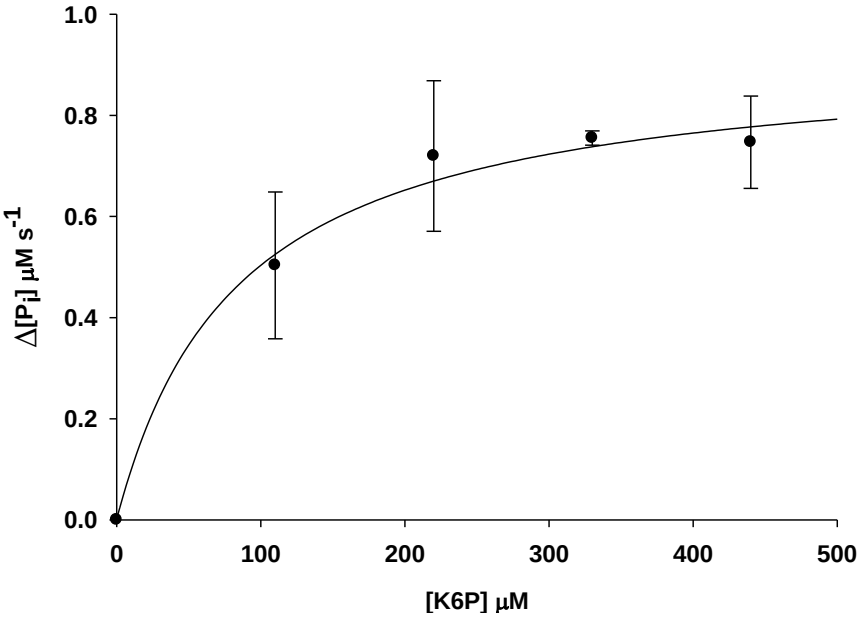
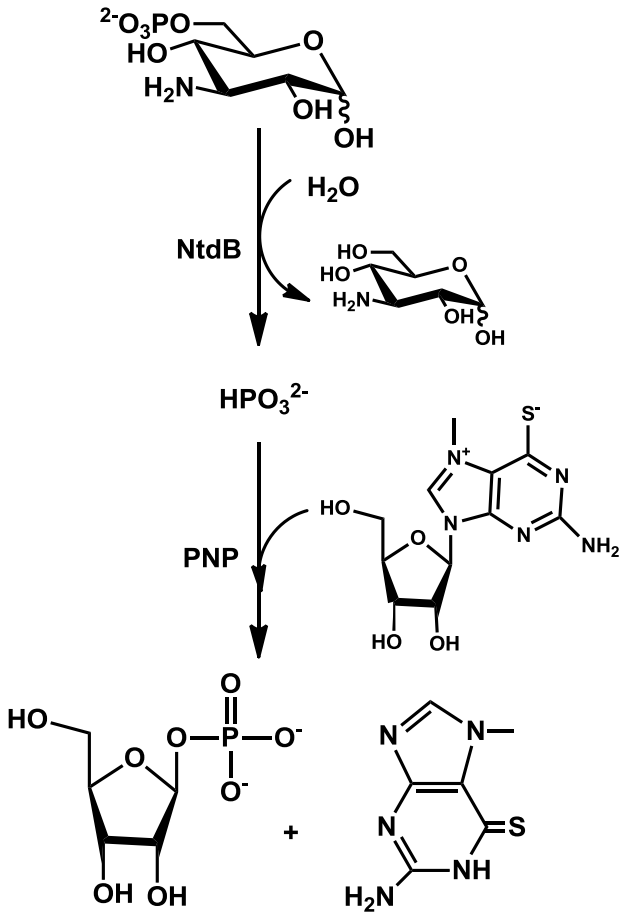


Figure 4.4: WT NtdB using malachite green assay.
Michaelis-Menten curve of WT NtdB generated by Sigma Plot 11.0.Each point represents at least three measurements

4.3.3 EnzChek® Phosphate assay

Twenty years ago, Webb proposed a coupled assay, in which two reactions take place.[80] The first reaction consists of the phosphatase under investigation dephosphorylating its substrate. The P_i released is used in the second reaction by purine-nucleoside phosphorylase (PNP) in the presence of 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG), producing two compounds, ribose-1-phosphate (R1P) and 2-amino-6-mercapto-7-methylpurine (AMMP). The latter product generates an absorbance peak at 360 nm which can be measured by UV spectrophotometry (**Scheme 4.2**). This is a continuous assay, since the release of P_i can be monitored as the reaction proceeds.



Scheme 4.2: Coupled reactions of the EnzChek® phosphate assay.
K6P is converted to kanosamine and P_i by the action of NtdB. The P_i is used to convert MESG to AMMP and R1P. The wavelength of AMMP was measured at 360 nm.

The EnzChek® phosphate assay is known to be effective in a narrow pH range and requires Mg^{2+} for activity. Therefore, the pH and metal ion-dependance of NtdB were not investigated.

4.3.4 Kinetic values of WT NtdB

The activity of an enzyme can be quantified using kinetics. The knowledge of the rate of an enzymatic reaction can help to predict a mechanism, identify substrates, and understand the function of an enzyme. The Michaelis-Menten equation (**3.2**) is a simple model establishing the relationship between the reaction rate and the concentration of

substrate for a fixed amount of enzyme. This model was used to establish the kinetics of NtdB and took into consideration its limitations: the amount of enzyme was much smaller than the amount of substrate used and only the initial rate of the reaction was measured.

Using the Enzchek® phosphate assay, NtdB was added to the reaction mixture containing various amounts of K6P. The formation of AMMP was monitored at 360 nm and the initial velocities were plotted against the substrate concentration. The initial velocities consisted of the straight-line portion of the graph, typically the first 60 seconds of the reaction. Data were fitted to the Michaelis-Menten equation using non-linear regression (**Figure 4.5**).

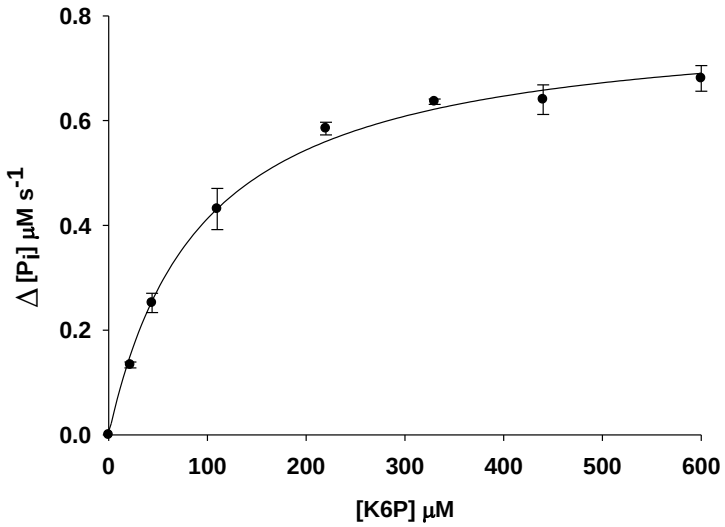


Figure 4.5: Hydrolysis of K6P catalyzed by WT NtdB.
Michaelis-Menten curve of WT NtdB generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

The kinetic constants obtained give us information about the behavior of the enzyme. The k_{cat} is a value indicating the turnover of substrate by an enzyme. The K_m relates to the dissociation constant of the enzyme for its substrate. The ratio of these values, k_{cat}/K_m, is therefore a measure of the efficiency of the enzyme, mostly known as specificity constant.

If an enzymatic or chemical reaction is completely efficient, where every encounter leads to a reaction, the maximum value of the specificity constant will be in the order of 10^8 to 10^9 $M^{-1}s^{-1}$, where the rate limiting step is the diffusion rate and not the rate of the reaction.[89] Most enzymes, in the presence of their native substrate, will present a specificity constant in the order of 10^5 to 10^8 $M^{-1}s^{-1}$.[90]

The kinetic values were calculated using three methods: non-linear regression, Lineweaver-Burk plot, and Eadie-Hofstee diagram. These two last plots were very popular when computer analysis was not accessible to all; researchers needed to rely on graphical representation of the data.

The Lineweaver-Burk is a double reciprocal plot based on the Michaelis-Menten equation yielding a straight line with the following equation (**4.1**):

$$\frac{1}{V_0} = \frac{K_m}{V_{max}}\left(\frac{1}{[S]}\right) + \frac{1}{V_{max}} \qquad \textbf{(4.1)}$$

The y-intercept corresponds to the inverse of the V_{max}, while the x-intercept represents the negative inverse of K_m (**Figure 4.6**). This plot does not give equal weight to all data points when calculating the slope, such that data obtained at small concentrations of substrate have a significant impact on the value of the slope. But this graphical representation is a rapid diagnostic tool to evaluate the fitness of the data obtained.

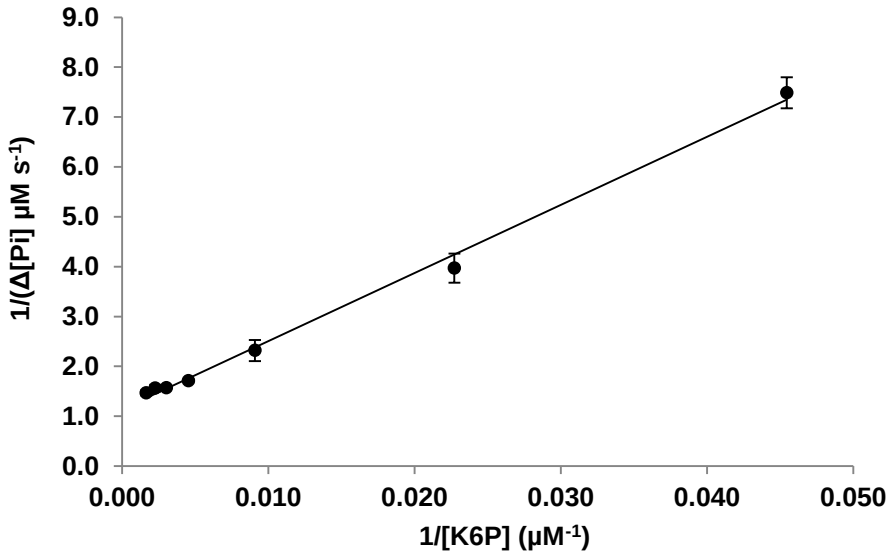


Figure 4.6: Graphical representation of the Michaelis-Menten equation using a Lineweaver-Burk plot.
The concentration of NtdB was 25 nM.

The Eadie-Hofstee diagram is another graphical representation of the Michaelis-Menten equation, where the reaction velocity is plotted against the ratio of velocity and substrate concentration (**4.2**).

$$V_0 = -K_m \left(\frac{V_0}{[S]}\right) + V_{max} \qquad \textbf{(4.2)}$$

The y-intercept represents V_{max}, while the slope of this graph represents the negative K_m (**Figure 4.7**). With this method, experimental errors are counted twice, since both axes are dependent on the initial velocities obtained and as such, neither are independent. But, it gives equal weight to all data points and is a rapid diagnostic tool. **Table 4.2** summarize the different kinetic values obtained with all three methods.

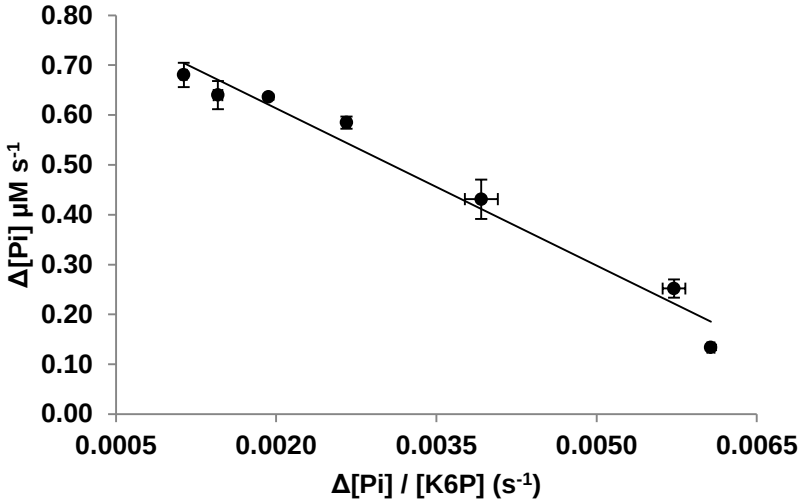


Figure 4.7: Graphical representation of the Michaelis-Menten equation using an Eadie-Hofstee diagram.
The concentration of NtdB was 25 nM.

Non-linear regression analysis resulted in the least amount of deviation between the calculated and the experimental data, and was therefore used to calculate kinetic values throughout this thesis.

Table 4.2: Kinetic values of NtdB obtained using different methods of data analysis.

Method used	V_{max} (µM s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat} / K_m ($M^{-1}s^{-1}$)
Non-linear regression	0.80 ± 0.02	0.093 ± 0.007	32 ± 3	(3.4 ± 0.4) x 10^5
Lineweaver-Burk	0.88 ± 0.15	0.120 ± 0.020	35 ± 15	(2.9 ± 1.0) x 10^5
Eadie-Hofstee	0.82 ± 0.11	0.105 ± 0.016	33 ± 6	(3.1 ± 0.7) x 10^5

In the literature, many members of the HAD superfamily have been kinetically investigated, but the comparison of interest lies mainly with other phosphatases belonging to the Cof family which possess a similar cap as NtdB. A phosphoglycolate phosphatase with high substrate specificity and a sugar phosphatase with a broad range of substrates have been characterized previously (**Table 4.3**). The highly specific phosphoglycolate phosphatase has a specificity constant in the order of 10^5 M^{-1} s^{-1}, similar to that of NtdB, while the non-specific sugar phosphatase is several orders of magnitude lower. This would suggest that the native substrate of NtdB is K6P.

Table 4.3: Comparing NtdB specificity with other members of the HAD superfamily.

PDB	Name	Substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat} / K_m ($M^{-1}s^{-1}$)
3GYG	NtdB	K6P	32 ± 3	0.093 ± 0.007	(3.4 ± 0.4) x 10^5
1L6R	Phosphoglycolate phosphatase[64]	Phosphoglycolate	8.3 ± 1.5	0.037 ± 0.003	2.2 x 10^5
1YMQ	Sugar phosphatase (BT4131)[65]	2-deoxy-D-glucose 6-phosphate	26 ± 1	3.9 ± 0.4	6.7 x 10^3

4.4 Putative substrates of NtdB

The activity of NtdB in the presence of K6P does not eliminate the possibility that it may be active with other substrates. Many enzymes are promiscuous or multispecific.[90] An array of substrates (**Figure 4.8**) was assayed with NtdB.

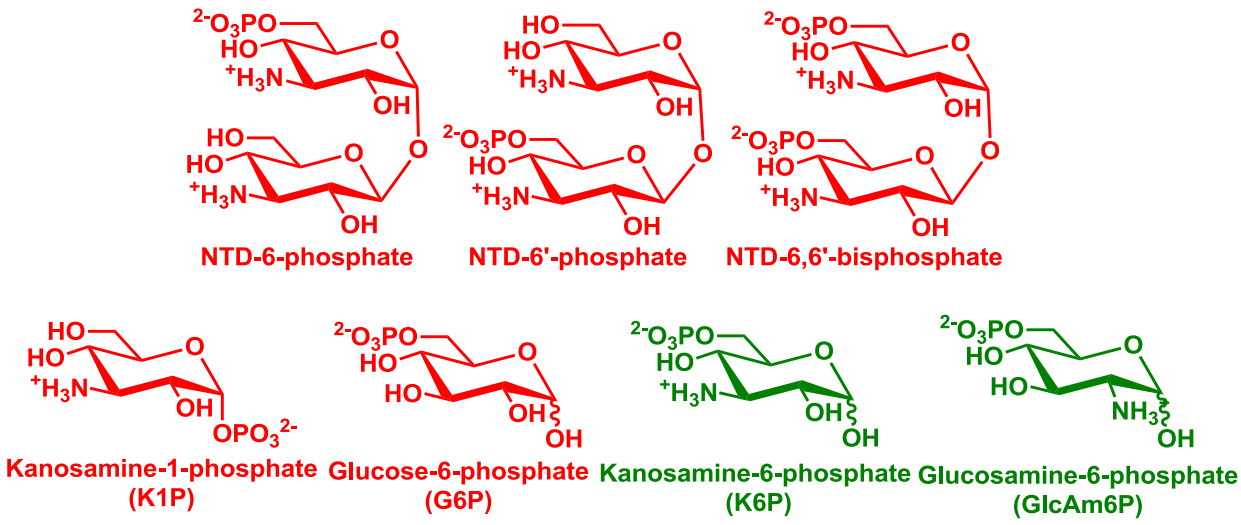


Figure 4.8: Different substrates tested for NtdB activity.

The structural similarity between trehalose-6-phosphate phosphatase (TPP) and NtdB (described in the introduction) highlighted the possibility that NtdB could be using a phosphorylated NTD as a substrate (**Scheme 1.4b**). To test whether this was the case, NTD-6-phosphate, NTD-6'-phosphate, and NTD-6,6'-bisphosphate were tested in the presence of NtdB. No phosphate release was detected for NTD-6,6'-bisphosphate. For

NTD-6-phosphate and NTD-6'-phosphate, very small but detectable activity was observed in the first minute of the reaction. There is doubt that this activity comes from the reaction of NtdB with the phosphorylated NTD because the activity disappears less than two minutes later, generating only a small amount of phosphate. The activity could be explained by some degradation of the phosphorylated NTD to K6P, or trace in the sample (contamination). NTD-6-phosphate and NTD-6'-phosphate are not commercially available and synthesis of more phosphorylated NTD would be required for further investigation, but the low level of phosphatase activity observed suggests that these are not good substrates of NtdB.

A substrate closely related to K6P is D-glucosamine-6-phosphate (GlcAm6P, **Figure 4.8**). The structural difference between them lies in the location of the amino group, at C3 for K6P and at C2 for GlcAm6P. A small amount of activity was detected from NtdB in the presence of GlcAm6P, but it was too small to be kinetically investigated and as such was not considered a good substrate for NtdB.

D-Glucose-6-phosphate (G6P) and kanosamine-1-phosphate (K1P) were also tested as substrates and neither triggered any activity from NtdB. This highlights the importance of the locations of the amino and phosphate group for substrate recognition and dephosphorylation by NtdB. In parallel with the observations made with GlcAm6P, NtdB seems to be highly specific for K6P. In the future, more substrates need to be assayed to validate this hypothesis.

4.5 Mutagenesis of *ntdB*

Once a substrate is found for an enzyme, such as K6P for NtdB, different experiments can be performed to try to explain the enzyme mechanism. One way to do so is by site-directed mutagenesis, in which targeted amino acid residues are selectively replaced with other residues. Effects on the kinetics can be observed and compared with the WT. Based on comparisons with other phosphatases, study of the crystal structure, and the knowledge that K6P is the native substrate, 13 different amino acid residues were

selected for mutagenesis for a total of 16 mutants, 9 located in the core of the enzyme and 7 in the cap (§4.8 and §4.9).

To perform site-directed mutagenesis, nucleotides needed to be changed in the *ntdB* gene sequence. Primers containing the desired mutation were designed, using the software OligoCalc, such that the mutated nucleotide(s) were located at least 4 nucleotides away from either end of the primer, the GC content was between 40% and 60% whenever possible, the length was 20 to 40 base pairs, and the melting temperature was between 55 °C and 65 °C (**Table 3.3**).

Site-directed mutagenesis followed the KAPA HiFi^{TM} PCR kit protocol, using different annealing temperatures for each mutant. **Table 4.4** summarizes the successful annealing temperature for each mutant, which was found by creating a temperature gradient across the heating block during the annealing step of PCR. Amplification of the desired mutation was confirmed by agarose gel electrophoresis.

Table 4.4: Annealing temperatures for mutants of NtdB

Mutation	T_m (°C)	Annealing (°C)
D25N	60.8	59
D27A	64.5	63 - 65
D27N	62.2	59.5 - 60.5
T65A	64.3	60 - 63
T65S	64.4	62 - 64
R149K	60.8	59
H152F	63.5	61 - 63
C187A	62.0	60 - 62
D193N	63.0	62 - 63
D196N	63.0	61 - 63
D199N	63.0	63
D201N	64.5	65
K209E	64.4	64
K209R	64.4	63 - 64
D232A	64.4	62 - 63
D236A	60.3	58

Plasmids containing the desired mutation were transformed into *E. coli* competent cells and after culturing onto agar plates, several colonies were observed, indicating

successful mutagenesis. The presence of the desired mutation in the plasmids was further confirmed by nucleotide sequencing.

4.6 Purification of mutants NtdB

The purification process used for all mutants of NtdB was the same as that used for the WT (§4.1). **Figures 4.9** and **Figure 4.10** show the purity of mutants in the core and cap domains, respectively.

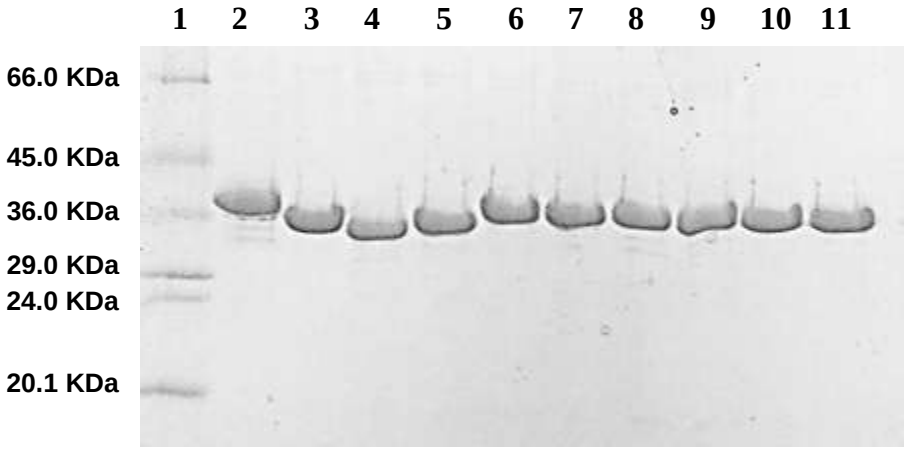


Figure 4.9: SDS-PAGE of pure WT NtdB and mutants in the core domain.
Lane 1: Ladder. Lane 2: WT NtdB. Lane 3: D25N. Lane 4: D27A. Lane 5: D27N. Lane 6: T65A. Lane 7: T65S. Lane 8: K209E. Lane 9: K209R. Lane 10: D232A. Lane 11: D236A.

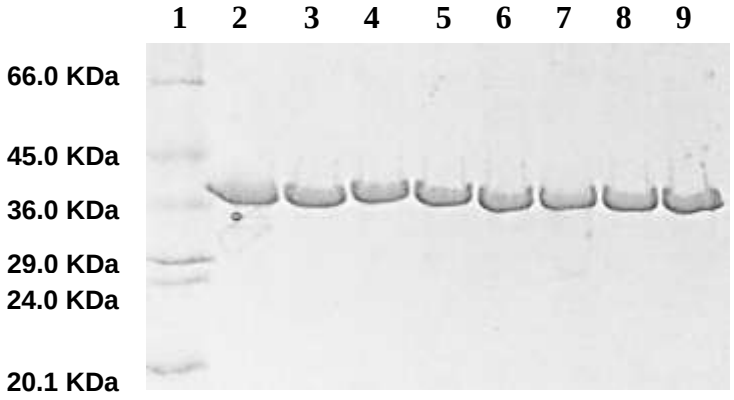


Figure 4.10: SDS-PAGE of pure WT NtdB and mutants in the cap domain.
Lane 1: Ladder. Lane 2: WT NtdB. Lane 3: R149K. Lane 4: H152F. Lane 5: C187A. Lane 6: D193N. Lane 7: D196N. Lane 8: D199N. Lane 9: D201N.

As was observed with the WT NtdB, the molecular weight of the mutants appeared higher than expected by SDS-PAGE, but still within error.[81, 82]

4.7 NtdB mutants of the core domain

Residues in the core domain of NtdB (D25, D27, T65, K209, D232, and D236) are believed to play a role in the proposed catalytic pathway of NtdB (**Scheme 1.6**), and belong to four motifs generally found in phosphatases.[91] Since almost all members of the HAD superfamily contain the same four conserved motifs in their core, the roles of these amino acid are believed to be the same in all members of the HAD superfamily. It was hypothesized that mutations performed in the core of NtdB would have similar effects to those performed in the core of other members of the HAD superfamily.

4.7.1 Motif 1

The first motif contains two important aspartate residues believed to be essential for activity; D25, a proposed nucleophile and D27, a proposed general acid/base.

The isosteric mutation of D25 to an asparagine residue was predicted to prevent activity of the enzyme. The neutral amid is much less nucleophilic than the negatively charged carboxylate, and this mutation would prevent the formation of the phospho-enzyme intermediate. D25N presented no activity in the presence of K6P, confirming the importance of this residue in the active site of the enzyme and giving credence to its role as a nucleophile (**Table 4.5**).

Table 4.5: Kinetic parameters of NtdB mutants of the core domain.

Motif	Mutation	K_m (μM)	k_{cat} (s^{-1})	k_{cat} / K_m ($M^{-1}s^{-1}$)	% activity
-	**WT**	93 ± 7	32 ± 3	$(3.4 \pm 0.4) \times 10^5$	100
1	**D25N**			No activity	
1	**D27A**			No activity	
1	**D27N**			No activity	
2	**T65A**			< 1% activity	
2	**T65S**	70 ± 6	22 ± 2	$(3.1 \pm 0.4) \times 10^5$	69 ± 6
3	**K209E**			No activity	
3	**K209R**			No activity	
4	**D232A**			No activity	
4	**D236A**			No activity	

D27 is believed to act as a general acid/base. This second aspartate residue was previously observed in the crystal structure of a phosphoserine phosphatase to be in a protonated state, forming two hydrogen bonds, one as acceptor with a neighboring threonine residue and one as donor to the phosphate.[62] D27A was predicted to be prevented such a role and the loss of activity of this mutant in the presence of K6P confirmed this prediction. The isosteric mutation, D27N, allows the formation of hydrogen bonds but prevents any role as a general acid/base. This mutant could have retained some activity through hydrogen bonding or proton-relay. But, the lack of activity observed for this mutant in the presence of K6P supports the idea that D27 is acting as a general acid/base (**Table 4.5**).

The results obtained for both aspartate residues in the first motif is in accordance with what has been observed in many other phosphatases. Site-directed mutagenesis performed on similar residues in other members of the HAD superfamily have yielded the same lack of activity (**Table 4.6**). Based on these results and the sequence alignment of NtdB with other members of the HAD superfamily, it was concluded that D25 and D27 very likely have the same roles as in other phosphatases, a nucleophile and a general acid/base, respectively.

Table 4.6: Comparing the percent activity of NtdB mutants in the core domain with other members of the HAD superfamily.

Motif	NtdB	%*	PSP[58]	%	PChP[92]	%	MDP-1[93]	%	sEHP[59]	%	HAD[94]	%
1	D25N	0	D20N	0			D10N	0	D9N	0	D10N	<1
			D20E	0	D31E	0	D10E	0	D9A	0	D10A	0
	D27A	0	D22E	0	D33E	0	D12N	8	D11A	0		
	D27N	0	D22N	0								
2	T65A	< 1	S109A	6	S166T	20	S68A	0.1	T123A	7	S118A	23
	T65S	69	S109T	115					T123N	3		
3	K209R	0	K158R	1	K242R	57	K99R	1.4	K160R	0	K151R	0.02
4	D232A	0	D179N	0.6			D122N	0.1	D185A	0		
			D179E	78	D262E	7						
	D236A	0	D183N	< 0.4					N189A	0	D180N	0
			D183E	63	D267E	2	N126D	52	N189D	0	D180E	<1

Abbreviations: PSP, phosphoserine phosphatase; PChP, phosphorylcholine phosphatase; MDP-1, Magnesium-dependant acid phosphatase; sEHP, soluble epoxide hydrolase phosphatase; HAD, haloacid dehalogenase.

* % = % activity relative to WT enzyme.

4.7.2 Motif 2

In NtdB, the second motif contains a threonine and its hydroxyl group was proposed to perform two roles: help orient the phosphate for the nucleophilic attack and stabilize the intermediate. This residue was mutated to an alanine, as well as a serine, to evaluate its importance.

T65 was mutated to alanine and was predicted to show no activity in the presence of K6P. This was the case and the lack of activity from the mutant is an indicator of the importance of the amino acid replaced. The T65A mutant showed less than 1% activity in the presence of K6P, underlining the importance of this residue for the catalytic activity of NtdB. This result is in accordance with similar mutations performed in other phosphatases, where, in most cases, removing the hydroxyl residue causes a considerable

reduction in the activity of the enzyme since it is preventing the formation of hydrogen bonds (**Table 4.6**).

The second mutation performed was T65S. Both serine and threonine residues contain a β-hydroxyl group to help stabilize the intermediate through hydrogen bonding but threonine also contains a methyl group, making this amino acid bulkier. Kinetic assays revealed that this mutant had reduced activity, but similar specificity (**Figure 4.11** and **Table 4.5**). The difference in size between both residues may explain why threonine is a better amino acid at this position; the bulkiness of threonine helps achieve a proper orientation of the phosphate group in the active site. But, the behavior of this amino acid in motif 2 varies between the members of the HAD superfamily. The mutation of this amino acid residue had caused a considerable decrease in activity in the case of phosphorylcholine phosphatase[92] and no change for phosphoserine phosphatase.[58] In both of these enzymes, a threonine was changed to a serine, opposite of NtdB (**Table 4.6**).

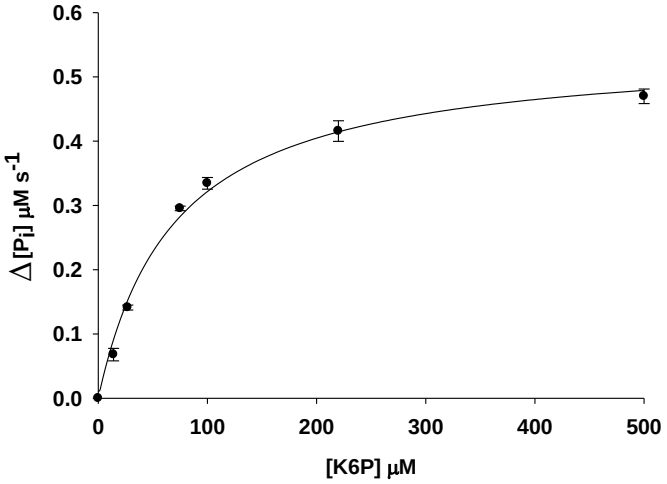


Figure 4.11: Hydrolysis of K6P catalyzed by T65S.
Michaelis-Menten curve of T65S generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

4.7.3 Motif 3

The third motif contains a conserved lysine residue which is believed to stabilize the intermediate and has been shown to form a hydrogen bond with the nucleophile in related phosphatases. This lysine residue was mutated to glutamate and arginine. As expected, K209E showed no activity in the presence of the substrate (**Table 4.5**). The negative charge introduced could have prevented the formation of a hydrogen bond with D25, and created repulsion with the substrate and/or intermediate, preventing any stabilization.[44, 64]

Since the mutant K209R conserved the positive charge and did not drastically modify the size of the residue, some activity should have been expected. However, no activity was observed in the presence of K6P for this mutant (**Table 4.5**). The effect of this mutation varies among phosphatase. In the case of phosphorylcholine phosphatase, the activity is only reduced approximately by half, still allowing the enzyme to function. Considerably reduced activity has been observed in others (**Table 4.6**). It is noteworthy that this lysine residue was reported to be poorly conserved in members of the HAD superfamily, compared to the other three motifs.[44]

From these results it is apparent that K209 is catalytically important. In the future, either a crystal structure of K209R or modeling of K209R in the closed-conformation of NtdB could help identify the impact of the mutation on the hydrogen bonds with the nucleophile and the intermediate.

4.7.4 Motif 4

Both aspartate residues of motif 4 are believed to help with coordination of the magnesium ion, positioning the metal, and increasing the electrophilicity of the phosphate group, favoring nucleophilic attack.[44, 62] The first aspartate would interact

directly with the magnesium ion while the second aspartate hydrogen bonds with a water molecule, which in turn coordinates with the magnesium ion.[53, 60, 61]

Two amino acid residues, D232 and D236, were mutated to alanine in the fourth motif. In both cases, the mutation caused a loss of activity of NtdB in the presence of K6P (**Table 4.5**), as was expected based on mutations reported for other phosphatases, in which even the mutation to asparagine caused a considerable reduction in activity (**Table 4.6**).

For NtdB, analysis of the different chains of the deposited crystal structure[47] reveals two different positions for the magnesium ion, the main difference being the coordination of either D236 or a water molecule (**figure 4.12**). In chains A, B, and D, D236 forms a hydrogen bond with a water molecule which coordinates the magnesium ion. In chain C, D236 coordinates directly to the magnesium, which is buried in the active site, and the water molecule seems to be in the position which would be occupied by the phosphate group of bound substrate.

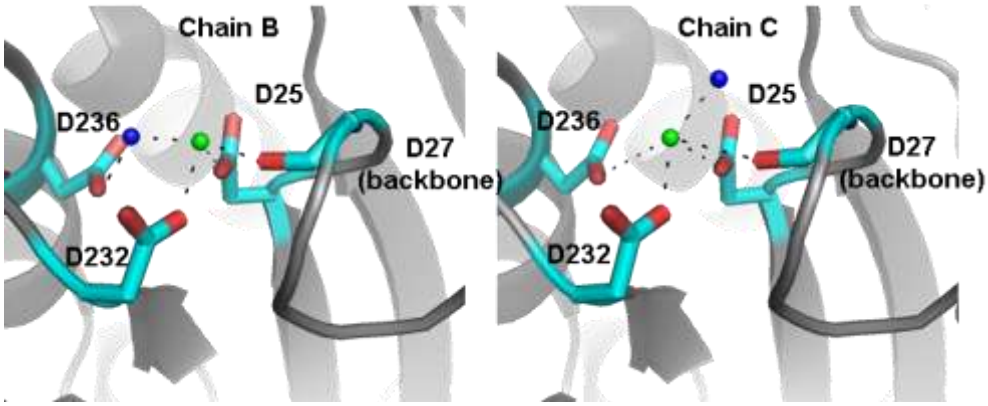


Figure 4.12: Comparing the magnesium ionic bonds of chains B and C of NtdB.
Green sphere represents magnesium and blue sphere represents a water molecule (PDB 3GYG).

A difference in the position of the magnesium ion was also observed in the crystal structure of a cyanobacterial sucrose phosphate phosphatase. The metal is coordinated through the water molecule in the open conformation. In the empty closed-conformation,

the metal is coordinated with the aspartate residue. When the crystals are soaked with the substrate sucrose 6-phosphate, both magnesium positions appear (**Figure 4.13**).[51]

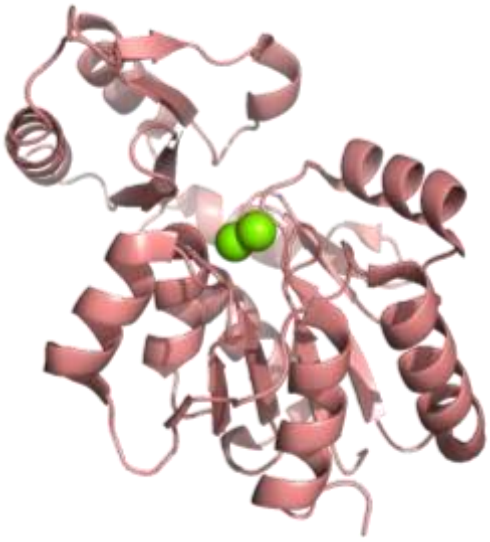


Figure 4.13: Two different positions for the magnesium ion in the closed-conformation of a sucrose phosphate phosphatase.
PDB 1TJ5

The following movement of magnesium ion is proposed for NtdB. In the open conformation of the enzyme, the magnesium favors the interaction with the water molecule. The phosphorylated substrate binds to the enzyme which then adopts a closed-conformation, moving the magnesium ion deep in the active site. The nucleophilic attack occurs and the enzyme re-adopt the open conformation, releasing the leaving group and giving solvent access to the active site. The magnesium moves back to its original position, where the hydroxide ion performs a nucleophilic attack on the phosphoenzyme intermediate and inorganic phosphate is released. This proposition is supported by the observations made on the crystal structures of sucrose-phosphatase in different conformations and with different substrates (PDB 1S20, 1TJ5, 2D2V, 2B1Q, and 1U2S).

The mutations performed on motif 1, 2 and 3 have put forth the importance of the alignment of the phosphate and the nucleophile for the enzyme to perform successfully. Similarily, the alignment of the magnesium, which coordinates the phosphate group,

would seem to be as important. Crystal structures of NtdB in the closed-conformation, with bound substrates, may help elucidate the movement of the magnesium ion.

4.8 NtdB mutants of the cap domain

As mentioned in the introduction, the cap is responsible for selectivity of each enzyme and differs among phosphatases. As more crystal structures of members of the HAD superfamily are solved, some commonalities among certain subcategories are starting to emerge. For example, phosphatases with a C2-type cap seem to possess a specificity loop, which orients some important residues of the cap.[64, 65] The specificity loop presents in NtdB comprises residues R186 to I203, and is longer than other sugar phosphatases.

A common methodology to initially identify important residues consists of modeling the desired substrate in the crystal structure of the enzyme, sometimes replacing a substrate already found in the active site. This method did not apply in this case since no crystal structure of NtdB in its closed-conformation exists at this time, and that modeling with the open conformation of NtdB using the software COOT[95] did not give any functional results.

Therefore, choosing amino acids in the cap of NtdB for mutagenesis was based on the observation that the location of the amino and phosphate groups of the substrate seem to be of primary importance for binding (§ 3.5). Focusing on the region of the specificity loop of NtdB, four aspartate residues were found in the cap domain: D193, D196, D199, and D201 (**Figure 4.14**). All four aspartic acids were mutated to asparagine, which removed the charge of these residues, but kept the size relatively constant. In the specificity loop, C187 was also identified near these aspartic residues and investigation of this amino acid through site-directed mutagenesis to alanine was performed. In the proposed substrate binding site formed by these four aspartate residues and the cysteine, the presence of two amino acids from a neighboring loop, R149 and H152, was observed,

and they were kinetically investigated. The pocket formed by these amino acid residues is the proposed substrate binding site.

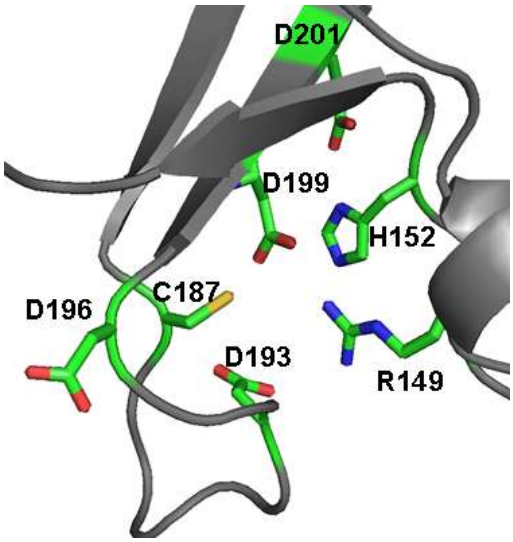


Figure 4.14: The amino acids of the cap domain surrounding the proposed substrate binding site.

4.8.1 D193N

D193 was initially believed to recognize the amino group of the substrate. This residue is located near the middle of the unusual long specificity loop. In the crystal structure of NtdB, this amino acid points in the direction of the active site. In the presence of K6P or GlcAm6P, D193N showed detectable but very small activity, and no activity with G6P (**Table 4.7**).

Table 4.7: Other substrates tested against NtdB and cap domain mutants.
(*** activity detected; * detectable but not quantifiable; - no activity).

Mutant	K6P	G6P	GlcAm6P
WT NtdB	***	-	*
R149K	***	-	*
H152F	***	-	*
C187A	***	-	*
D193N	*	-	*
D196N	***	-	*
D199N	-	*	*
D201N	***	*	-

The drastic negative effect of this mutation on the activity of NtdB with K6P would indicate the importance of this amino acid (**Table 4.8**). But, since there was no activity with G6P, in which the only difference is the presence of a hydroxyl group instead of an amino group, this residue may play a role other than substrate recognition. Sometimes, the cap also contains residues that play a role either in substrate turnover or as a structural component.[44] Due to its location, it is suggested that D193N plays a structural role in keeping the enzyme in the closed-conformation. This hypothesis would need to be further explored crystal structures of NtdB in its closed-conformation.

Table 4.8: Kinetic parameters of NtdB mutants of the cap domain.

Mutation	K_m (μM)	k_{cat} (s^{-1})	k_{cat} / K_m ($M^{-1}s^{-1}$)	% activity
WT	93 ± 7	32 ± 3	(3.4 ± 0.4) x 10^5	100
R149K	92 ± 9	38 ± 4	(4.1 ± 0.6) x 10^5	120 ± 12
H152F	110 ± 10	44 ± 4	(4.0 ± 0.5) x 10^5	140 ± 13
C187A	333 ± 46	0.58 ±0.05	(1.7 ± 0.3) x 10^3	1.8 ± 0.2
D193N	Less than 1% activity			
D196N	134 ± 13	26 ± 3	(2.0 ± 0.3) x 10^5	81 ± 9
D199N	No activity			
D201N	96 ± 9	1.6 ± 0.2	(1.7 ± 0.3) x 10^4	5.0 ± 0.6

4.8.2 D196N

D196 is located before the last anti-parallel β-strand of the cap and points out of the active site. Because of its position, the effect of mutating this aspartic acid was expected to be low. As seen in **Table 4.8** and **Figure 4.15**, the activity was reduced to 80% and its specificity reduced by less than half. It also has the same behavior as the WT NtdB in the presence of other substrates, such that there is no activity with G6P and some detectable, but very small, activity with GlcAm6P (**Table 4.7**). This would indicate that this residue is not a major player in the activity of the enzyme, or that its role is to form hydrogen bonds with one or more hydroxyl groups of K6P, hydrogen bonds that would still be present when mutated to asparagine.

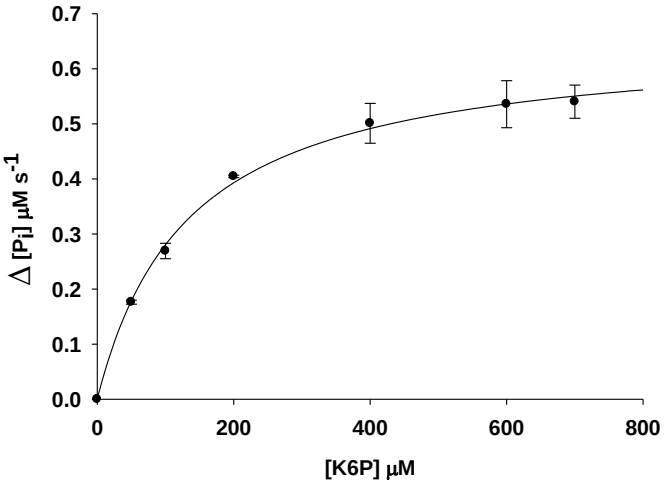


Figure 4.15: Hydrolysis of K6P catalyzed by D196N.
Michaelis-Menten curve of D196N generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements..

The crystal structure of a sucrose phosphate phosphatase with secondary structure similar to NtdB was solved in both the open and closed-conformation, allowing a movement of the cap to be observed (**Figure 4.16**).[51] Because the specificity loop of NtdB is seven amino acids longer than the specificity loop of this sucrose phosphate phosphatase, movement of the cap may cause a change in orientation of D196, allowing

for contact with K6P. This hypothesis could be confirmed with a crystal structure of the closed-conformation of NtdB.

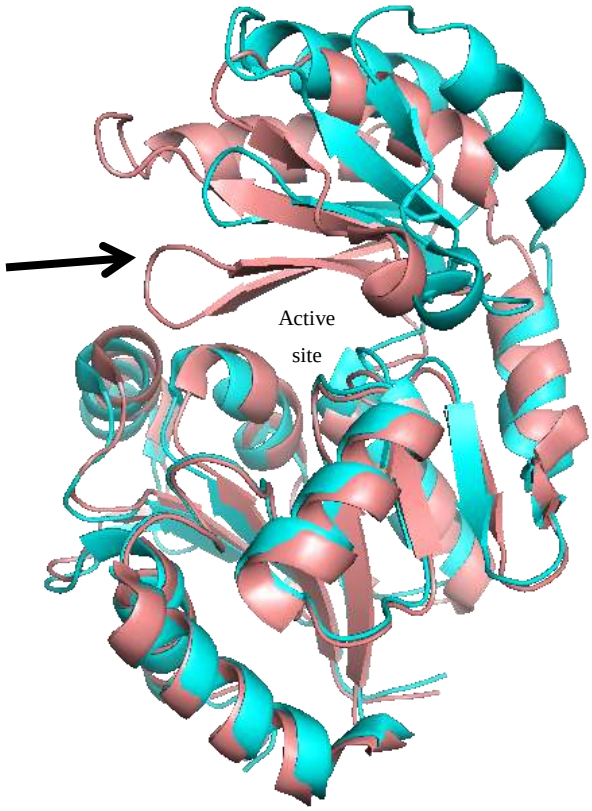


Figure 4.16: Movement of the cap between the open conformation and the closed-conformation of a sucrose phosphate phosphatase.
The arrow indicates the specificity loop in the closed-conformation. Open: Cyan, PDB 1S20; Close: Salmon, PDB 1TJ5.

4.8.3 D199N and D201N

D199 was proposed to play a central role in the binding of the substrate in the cap domain. The negatively charged carboxylate was presumed a key feature in the recognition of the amino group. The mutation to asparagine removed this charge, while keeping its potential for hydrogen bonds. In the presence of K6P, D199N showed no activity (**Table 4.8**), while in the presence of G6P, there was a small amount of activity that was detected (**Table 4.7**). This change in activity with both substrates supports the proposition that this amino acid is the main residue responsible for the recognition of the

amino group of K6P by NtdB. This would also be consistent with the proposed substrate binding site. Further investigations of this amino acid residue are proposed, such as mutation to alanine to increase the substrate binding pocket size and remove any hydrogen bonding capability.

D201 is proposed to have a similar role to D199. When mutated to asparagine, some changes in the substrate specificity of NtdB occur. With K6P, reduction of activity to 5% was observed (**Figure 4.17**). Additionally, no more activity was observed with GlcAm6P, while some activity was now seen with G6P (**Table 4.7**).

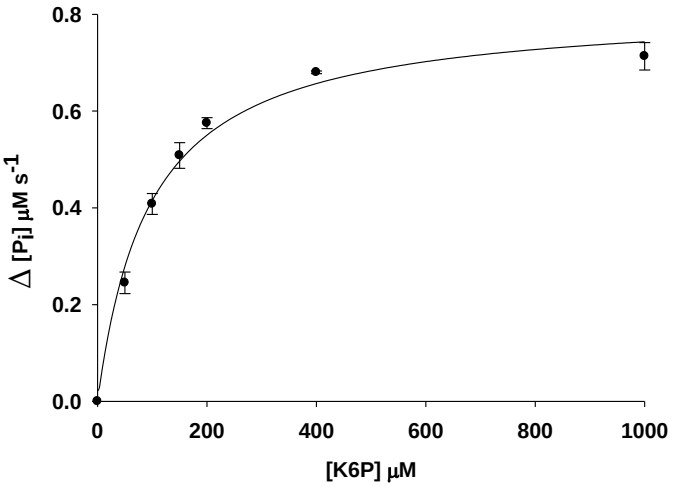


Figure 4.17: Hydrolysis of K6P catalyzed by D201N.
Michaelis-Menten curve of D201N generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

This change in substrate specificity not only indicates that this amino acid is an integral part of the substrate binding site, but also helps in proposing an orientation for K6P in the binding site. **Figure 4.18a** illustrates the binding interactions of the amino group of K6P to residues D199 and D201. GlcAm6P is characterized by an amino group at C2 instead of C3, and D201 may be the only residue holding this amino group in place (**Figure 4.18b)**. Mutation to D201N would prevent the stabilization of the amino group and the enzyme would not recognize this substrate anymore.

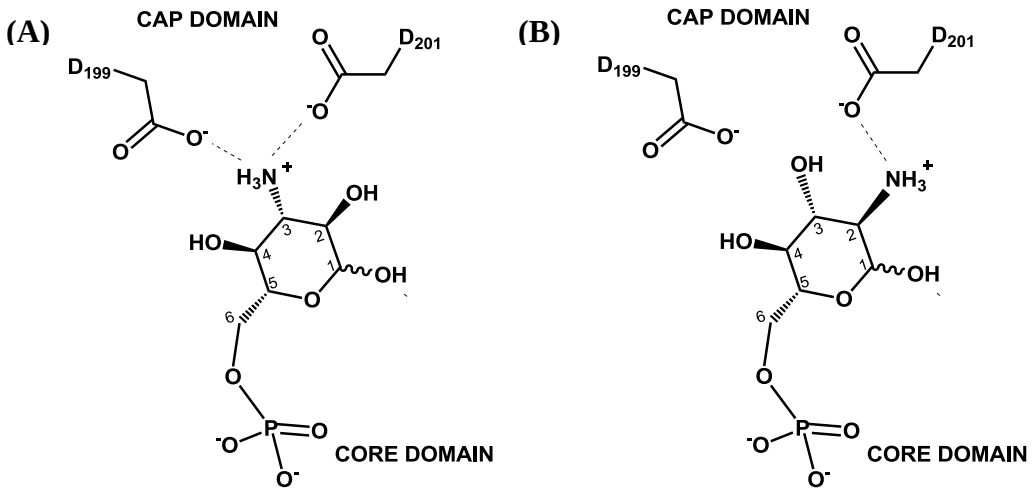


Figure 4.18: Proposed orientation of the substrate with respect to D199 and D201.
(A) K6P; (B) GlcAm6P

It is proposed that D199 and D201 represent the main amino acid residues responsible for the binding of the substrate. Both are on a β-strand and their relative orientation is not expected to change much upon movement of the cap domain. Their negative charge attracts the amino group of K6P and orients the substrate in the right position, allowing the enzyme to adopt the closed-conformation, and the nucleophile to attack the phosphate group. This combined negative charge repels the hydroxyl group of G6P and removal of either negative charge allows for small but detectable activity to occur with G6P. Again, a crystal structure of NtdB in its closed-conformation would help to confirm this hypothesis.

4.8.4 C187A

Due to its location (**Figure 4.19**), it is proposed that the role of C187 is to orient D193 and D199. The mutant C187A would not be able to form the hydrogen bonds necessary to orient the aspartate residues properly.

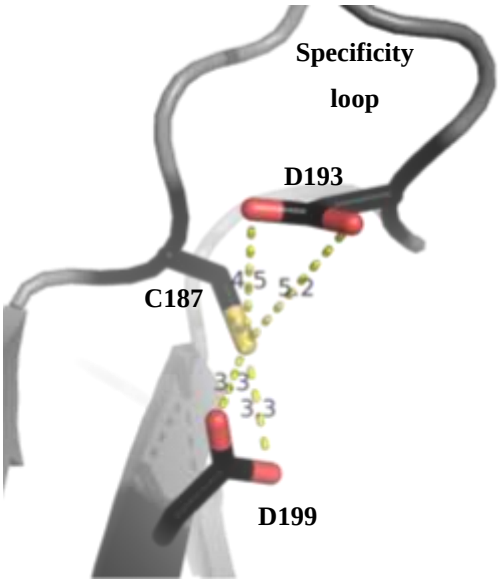


Figure 4.19: Hydrogen bond distance of C187 to D193 and D199 in the open conformation of NtdB.

The mutation C187A reduced the value of k_{cat} with K6P to 2%, while the K_m is more than three times higher than with the WT (**Figure 4.20** and **Table 4.8**). This would indicate that binding of the substrate is impaired, which could hinder and slow down the reaction..

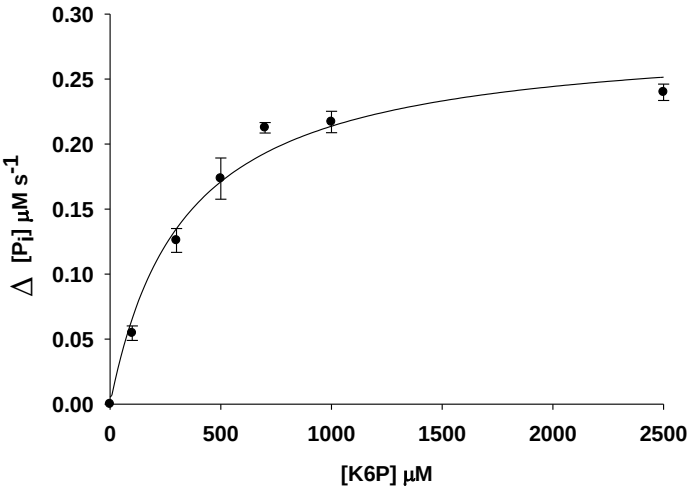


Figure 4.20: Hydrolysis of K6P catalyzed by C187A.
Michaelis-Menten curve of C187A generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

Since D199 is likely responsible for binding the amino group and orienting the substrate, the lack of hydrogen bonds to D199 would make it more difficult for the substrate to bind, explaining the increase in $K_{m.}$ C187 is also near D193, and the movement of the cap when the enzyme is adopting its closed-conformation may cause hydrogen bonds to form between both residues, helping to orient D193. If, as proposed, D193 is responsible for the proper structure of the closed-conformation, the lack of hydrogen bonding by C187A would impair orientation of D193, and prevent the reaction from proceeding, causing the drastic reduction observed in k_{cat}.

Therefore, it is proposed that C187 would play more of a structural role in the substrate binding site by helping to orient these two important aspartate residues.

4.8.5 R149K and H152F

In the open conformation of NtdB, two other residues from a neighboring loop, R149 and H152, are observed in the region of the substrate binding site. It is proposed that in the open conformation, these two amino acid residues are positively charged and

help stabilize the negative charges of D193, D199, and D201 (**Figure 4.21**). In the closed-conformation, the neighboring loop to which they belong would be moved away from the substrate binding site.

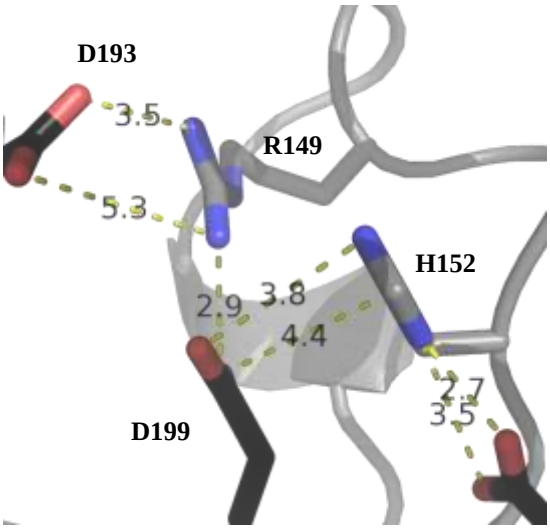


Figure 4.21: The hydrogen bond distance from R149 and H152 to aspartate residues in the open conformation for NtdB.

R149K, which has a reduced side-chain length while conserving its positive charge, caused no change in activity, the values within error of the WT NtdB (**Figure 4.22** and **Table 4.8**). The positive charge may be the important feature of residues at this position but further mutations would be required to verify the importance of this residue in the substrate binding site.

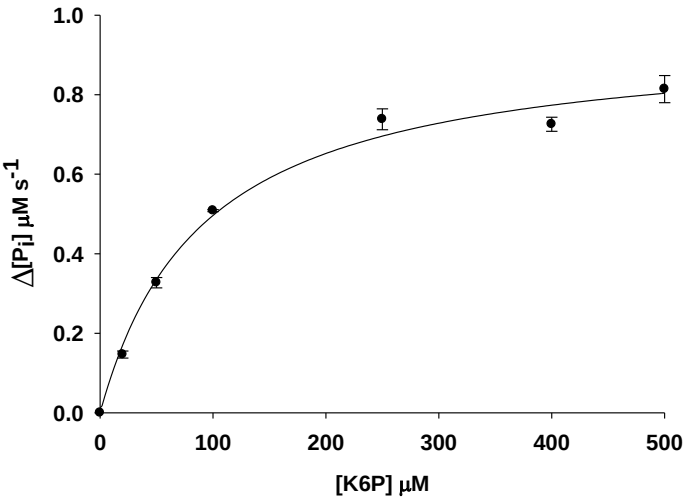


Figure 4.22: Hydrolysis of K6P catalyzed by R149K.
Michaelis-Menten curve of R149K generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

Histidine residues have a pK_a of 6 and are not expected to be positively charged at neutral pH. But, enzymes create microenvironments which may locally affect the pH and as such H152 may be positively charged. H152F removed this positive charge and any potential hydrogen bonding, while keeping the aromatic properties.[96] An increase of activity up to 140% was observed (**Figure 4.23** and **Table 4.8**) and may be due to a better access of the substrate to the aspartate residue in the binding site. These results would suggest that H152 is positively charged and that the charge hinders the activity of the enzyme. As with the other mutations in the cap, the crystal structure of the closed-conformation of NtdB would validate this assumption.

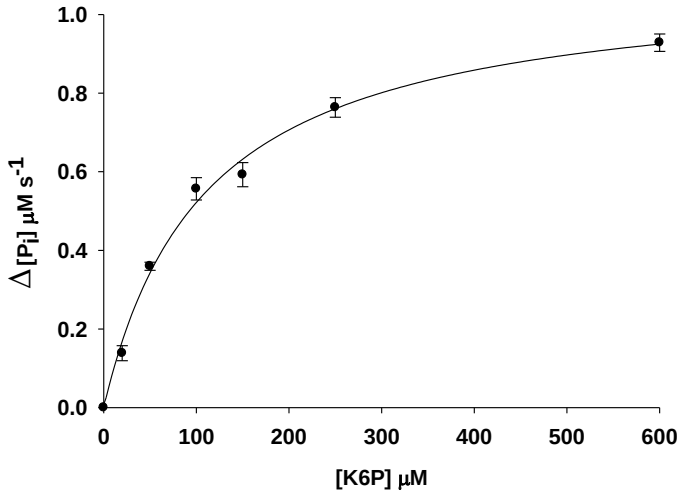


Figure 4.23: Hydrolysis of K6P catalyzed by H152F.
Michaelis-Menten curve of H152F generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

4.9 Prevalence of NtdB-like enzymes

NtdB's unique feature compared to other phosphatases in the HAD superfamily is the presence of a long specificity loop in the cap domain, which contains amino acid residues that are important for the binding of K6P. A search in GenBank® was performed to evaluate whether or not this specificity loop was common in phosphatases of other organisms (**Figure 4.24**).[68]

This research has revealed that in bacteria, many amino acid sequences from putative phosphatases have been found containing this specificity loop (**Figure 4.24**). Of these, the overall amino acid sequences of members of the *Bacillus* genus are more than 70% similar compared to NtdB from *Bacillus subtilis*, while other organisms are more than 50% similar, regardless of the phylum they belong to. K6P hydrolysis may be a common phenomenon in the bacterial kingdom.

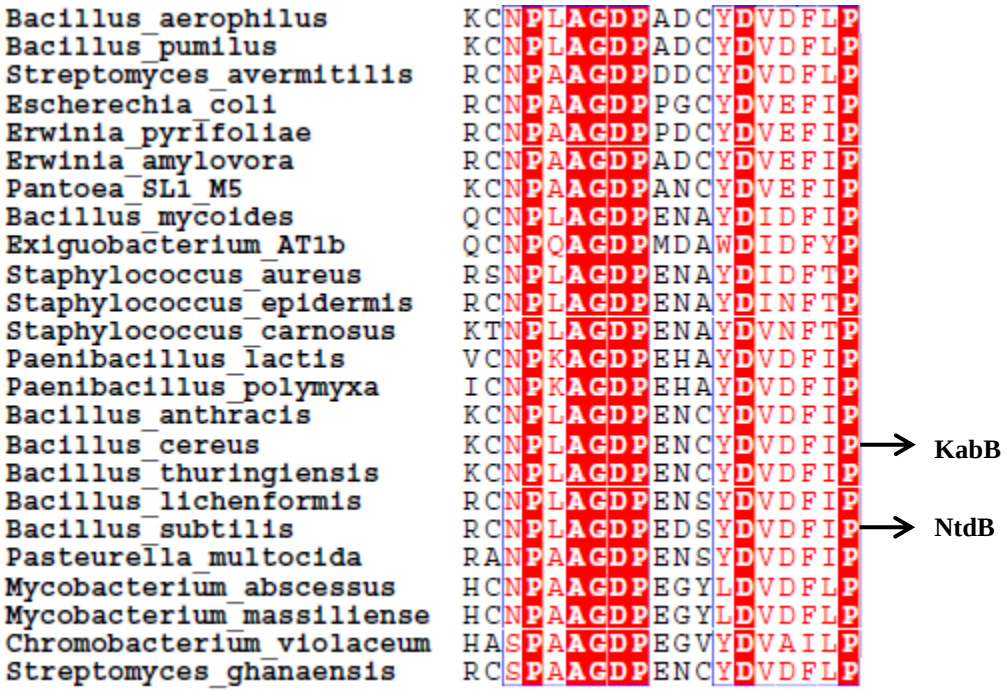


Figure 4.24: Sequence alignment of the specificity loop of NtdB with other proteins.
Generated with *ClustalW*[20], using the Blosum scoring matrix[21]: opening gap penality = 10; extending gap penality = 0.20; distance gap penality = 5. Image was created by *ESPript* V2.2[22]. The amino acids represented with white letters are identical, while red letters denote similar residues.

As can be seen in **Figure 4.24**, two aspartic residues are strictly conserved among all species and correspond to D193 and D199 in NtdB. In the last section, these amino acids were shown to be necessary for the enzyme to recognize K6P as a substrate. D199 was proposed to play a central role in substrate recognition, while D193 may be a key player in the closed-conformation of the enzyme. It is proposed that the presence of this long specificity loop in the cap domain, containing two conserved aspartic acid residues, is a feature of enzymes catalyzing the dephosphorylation of K6P. To test this hypothesis, KabB, which has been proposed to be involved in kanosamine biosynthesis,[23] was investigated.

4.10 Homology modeling of KabB

Comparative modeling allows for generation of a structural model of a target protein, based on the comparison with a similar protein of known structure. The high amino acid sequence similarity between NtdB and KabB suggests that both proteins have similar secondary structures. Using SWISS-MODEL[71-73], **Figure 4.25** shows no difference between the secondary structures of NtdB and KabB, which is in accordance with CD experiments (§4.14). The accuracy of this model was calculated to be 73% (as defined by equivalent Cα positions).

Figure 4.25: The comparison between NTD and the proposed crystal structure of KabB.
NtdB is represented in green (PDB 3GYG) and KabB in orange.

The active site residues in the core domain are similar in orientation between the proposed KabB model and NtdB (**Figure 4.26**). In the region of the substrate binding site located in the cap domain, the main differences in the orientation of the amino acids are R149 in NtdB, and K154 in KabB (**Figure 4.27**). In NtdB, R149K did not significantly affect the activity of the enzyme.

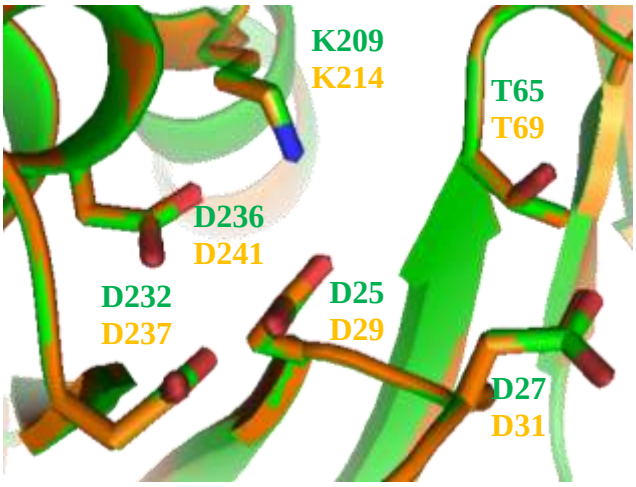


Figure 4.26: The comparison between NTD and the proposed location of the amino acid residue of KabB in the active site.
NtdB is represented in green (PDB 3GYG) and KabB in orange.

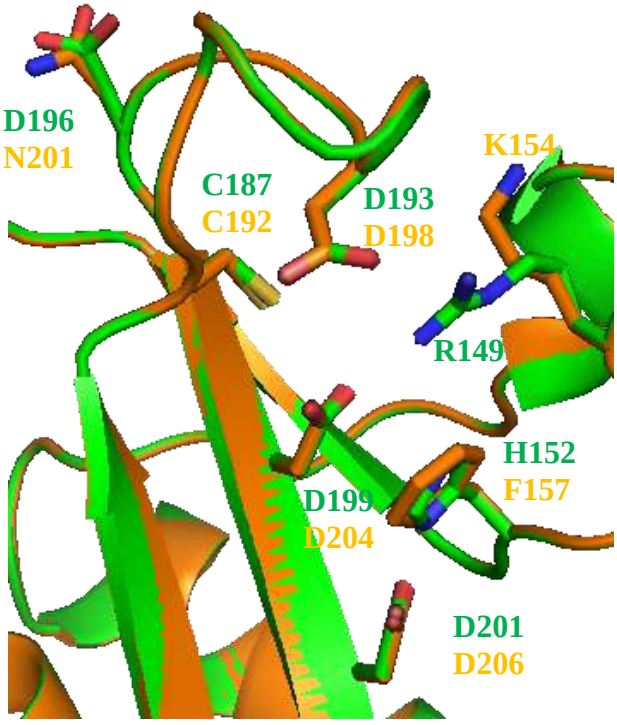


Figure 4.27: The comparison between NTD and the proposed location of the amino acid residue of KabB in the substrate binding site.
NtdB is represented in green (PDB 3GYG) and KabB in orange.

Homology modeling of KabB suggests that KabB possesses the structural characteristics of a phosphatase. In its core domain, a Rossmann-like fold is present, as

well as a β-hairpin. The cap domain not only resembles the Cof hydrolase family, with its three β-strands, but also contains the unusually long loop characteristic of NtdB. This model does not suggest any differences between the structure of NtdB and KabB.

4.11 Sub-cloning *kabB*

According to BLAST[45] and SWISS-MODEL,[71-73] NtdB contains an amino acid sequence identity of 55% with KabB (**Figure 1.3**) and it was believed that the comparison of the kinetic behavior of both enzymes with K6P would reveal a strong similarity. To perform such a comparison, the *kabB* gene was obtained from the kab operon in the pET-22b-*kabABC* construct.

First, the *kabB* gene needed to be isolated and subcloned into another vector. The vector pET-28b was selected as it was the same plasmid used for NtdB. Therefore, a similar construct from the same plasmid would be obtained, which allows for a direct comparison.

The restriction sites used to introduce the *kabB* gene in the vector were *Nde*I and *Bam*HI (**Figure 4.28**). DNA sequencing confirmed successful sub-cloning of *kabB*.

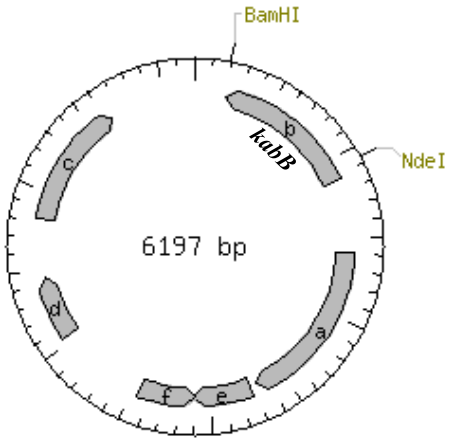


Figure 4.28: Vector map of *kabB* in pET-28b using unique restriction sites BamHI and NdeI.
Figure generated using NEBcutter v2.0.[70]

4.12 Purification WT KabB and mutant

The successful purification of WT KabB can be observed in **Figure 4.29**.

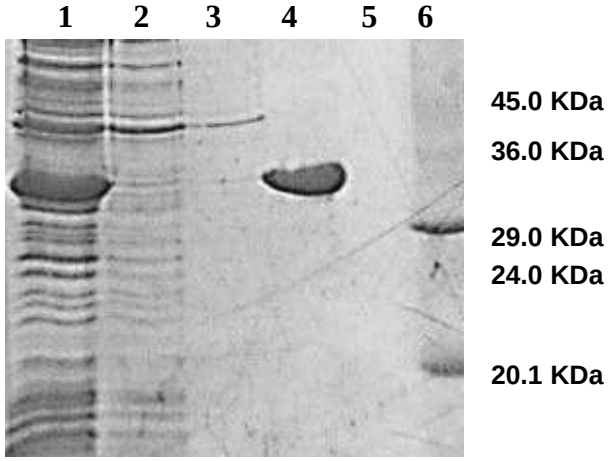


Figure 4.29: SDS-PAGE of pure WT KabB.
Lane 1: Crude WT KabB. Lane 2: Flow-through. Lane 3: Wash. Lane 4: Eluted WT KabB. Lane 5: Empty. Lane 6: Ladder.

The calculated molecular weight of KabB was 34 824 Da, and the apparent molecular weight on the SDS-PAGE reflects this value. The KabB N201D mutation was also successful and resulted in an apparent molecular weight similar to WT KabB (**Figure 4.30**).

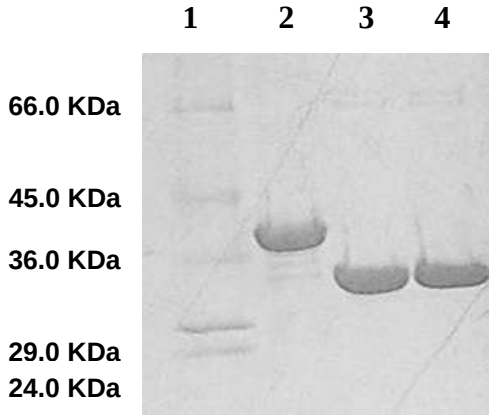


Figure 4.30: SDS-PAGE of pure KabB N201D.
Lane 1: Ladder. Lane 2: WT NtdB. Lane 3: WT KabB. Lane 4: N201D.

The difference between the apparent molecular weight of NtdB by SDS-PAGE and KabB is the first indication that KabB and NtdB, in spite of their high amino acid sequence identity, may have different characteristics. This was further observed in the kinetics with K6P.

4.13 Kinetics of KabB

Kinetics were performed on KabB using the EnzChek® phosphate assay (**Figure 4.31**) and compared with WT NtdB (**Table 4.9**).

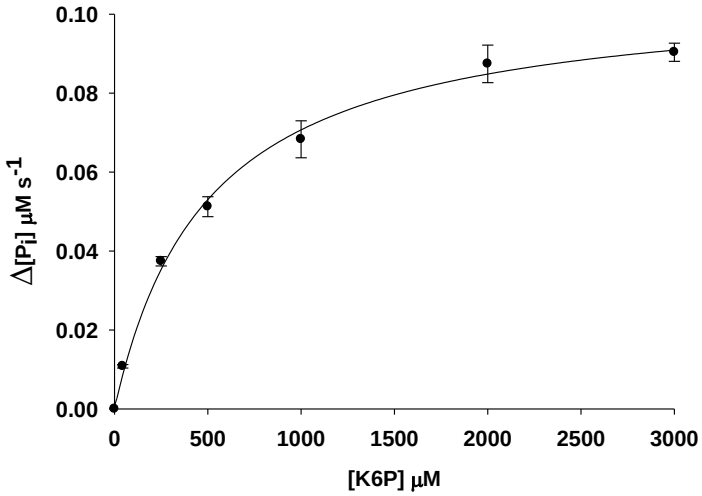


Figure 4.31: Hydrolysis of K6P catalyzed by WT KabB.
Michaelis-Menten curve of WT KabB generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

Table 4.9: Comparison of the kinetic values of NtdB, KabB, and KabB N201D.

Enzyme	V_{max} (nM s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat} / K_m ($M^{-1}s^{-1}$)
NtdB	800 ± 20	93 ± 7	32 ± 3	(3.4 ± 0.4) x 10^5
KabB	106 ± 3	500 ± 50	0.42 ± 0.04	(8.5 ± 1.0) x 10^2
N201D	142 ± 7	570 ± 60	0.57 ± 0.06	(8.8 ± 1.0) x 10^2

Because of the strong amino acid sequence similarity between KabB and NtdB, it was expected that both enzymes would present similar kinetic results with K6P. However

this was not the case. The K_m of KabB was 5 times higher than that of NtdB, while the activity of KabB was 1.3% of NtdB.

Analysis of specificity loop sequences of both enzymes (**Figures 1.3 and 4.24**, NtdB amino acids R186-I203, KabB residues K191-I208) reveals only three amino acid differences: K191, N201 and C202 (equivalent to R186, D196, and S197 in NtdB). The first amino acid difference, arginine/lysine contains the same positive charge and was not expected to create much impact since its location is far from the substrate binding site. The D196N mutation was also performed in NtdB (§4.8.2) and the change in kinetics were not comparable to what was observed with KabB. Nevertheless, the mutation of this amino acid was also performed in KabB, N201D, and results indicate that the mutation slightly increases the activity KabB with K6P, but the K_m still remains much higher than with NtdB (**Figure 4.32**). This confirms that the presence of an aspartic acid at this position of the substrate binding site of NtdB contributes to catalysis with K6P but still does not account for the notable difference in the kinetics between NtdB and KabB.

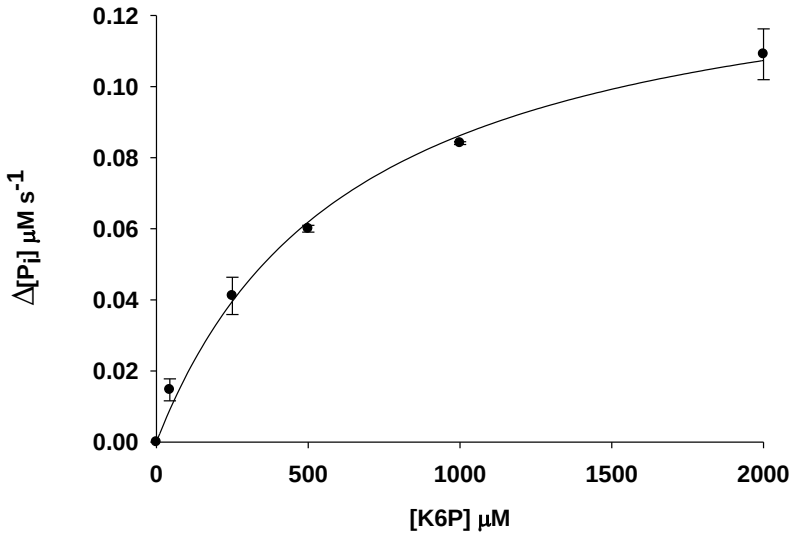


Figure 4.32: Hydrolysis of K6P catalyzed by KabB N201D.
Michaelis-Menten curve of N201D generated by Sigma Plot 11.0, using the EnzChek® Phosphate assay. Each point represents at least two measurements.

The third amino acid difference in the specificity loop corresponds to C202 in KabB, and S197 in NtdB. The presence of this cysteine may affect the substrate binding site of the enzyme if a disulfide bridge is formed with another cysteine residue. C192 in KabB (C187 in NtdB) may be a good candidate for the formation of a disulfide bond with C202. In the crystal structure of NtdB, the corresponding amino acids, S197 and C187, are 5.3 Å apart from each other (**Figure 4.33**) and seem to form a dihedral angle around 90°, which is an angle favored for the formation of disulfide bonds.[97] So, if S197 was replaced by a cysteine residue, a disulfide bond may occur. The mutation of C187 to alanine in NtdB did have a strong impact on the activity and the specificity of the enzyme (§4.8.4). The formation of a disulfide bond in KabB would prevent this residue from performing its proposed structural role. Further investigation of C202 in KabB, such as a mutation to serine, would be required in the future to test the hypothesis that this amino acid difference between KabB and NtdB contributes for the difference in activity seen with K6P.

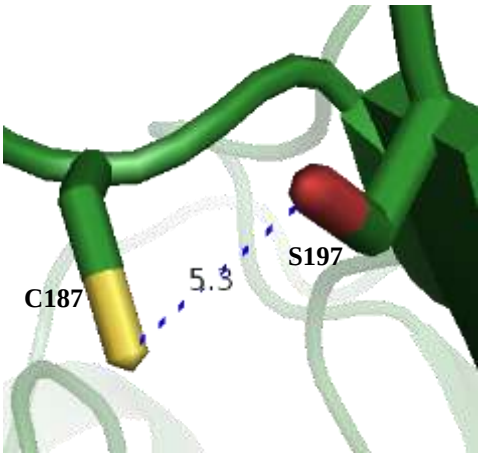


Figure 4.33: Distance between the cysteine residue and the serine residue in WT NtdB

On a neighboring loop, the mutation R149K in NtdB showed no change, while H152F increased the activity of the enzyme (**Table 4.7**). These amino acids correspond to K154 and F157 respectively in KabB. Therefore, these changes in the amino acid sequence could not be responsible for the difference observed in KabB. The difference from histidine to phenylalanine could have suggested that KabB would work even better

than NtdB in the presence of K6P, however this was not observed. Therefore, another factor must be interfering.

In the core domain, the four motifs are also strictly conserved in KabB, as well as most amino acids around them. If the lower activity of KabB compared to NtdB is not due to the presence of the cysteine residue C202 in the substrate specificity loop, it could be due to an improper closing of the protein. The crystal structure of KabB would then be needed to compare minor structural difference with NtdB.

It is possible that K6P is not the native substrate of KabB. To start testing this hypothesis, kinetics were performed with G6P, GlcAm6P, NTD-6'-phosphate and *para*-nitrophenylphosphate. None of these substrates showed detectable activity with KabB. Other substrates would need to be tested in the future, such as NDP-kanosamine.

4.14 Circular dichroism

Circular dichroism (CD) is a useful tool to determine if a mutation causes any significant changes in the structure of a protein. It uses the optical property of the backbone of proteins, which is absorbing left and right circularly polarized light in an unequal amount. Depending on the secondary structure of the protein, this gives rise to different spectra.[98,99] This electronic absorption property can be measured by CD and can be used to evaluate the changes in the secondary structure of the protein in different conditions.[100] Based on previous studies of proteins with unique secondary structure, α-helices are expected to produce negative peaks at 208 nm and 222 nm and a positive peak at 193 nm[101], β-antiparallel sheets show negative peak at 218 nm and positive peak at 195 nm[102], and disordered proteins contain a negative peak at 195 nm and positive peak at 210 nm.[103] A protein spectrum is therefore the sum of the peaks created by the amounts of different secondary structures.

Because some amino acids may be instrumental in the formation of an α-helix or β-strand, or in the prevention of such secondary structures, a mutation in certain amino acid

can cause the protein to precipitate, unfold or lose some of its structural properties, which can be measured by CD.

Since mutations were introduced into NtdB and KabB to evaluate the effect of selected amino acid residues on the activity of the enzyme and not on the secondary structure, CD spectra of NtdB and each mutant were collected and compared to the WT to verify the conservation of secondary structure. The results are normalized to account for difference in concentration as shown in **Figure 4.34**.

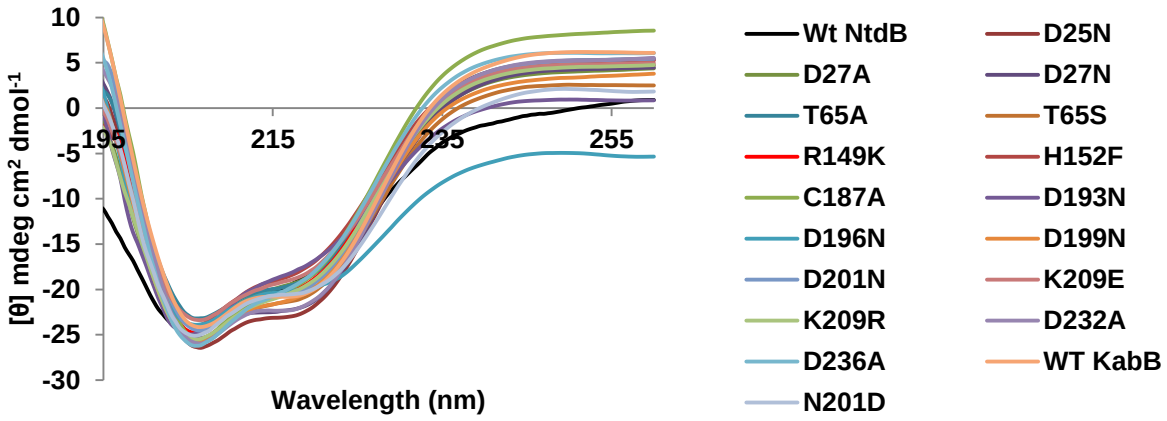


Figure 4.34: Comparing CD spectra of KabB, NtdB, and mutants.
[θ] corresponds to the molar ellipticity.

Comparing the format and shape of a curve can be subjective to the observer. Comparison of the spectrum can therefore be performed by analyzing the secondary structure of a protein by numerous software programs, which make use of database of known secondary structure of proteins. These programs deconvolute the CD spectra and calculate percentages of each secondary structure, which can then be compared. Comparing the validity of these softwares is beyond the scope of this thesis and since CDNN[77] was chosen by the SSSC at the University of Saskatchewan, in Saskatoon, it was selected for this thesis. **Table 4.10** shows the analysis obtained for the WT NtdB and its mutants using a simple basis set between 200-260 nm.

Table 4.10: Comparison of the percent content of secondary structures for NtdB and its mutants

Mutant	Concentration (mg/mL)	α-helices (%)	β-antiparallel (%)	β-parallel (%)	β-turn (%)	Random (%)	Total (%)
WT NtdB	**0.46**	**29**	**11**	**9.2**	**18**	**31**	**98**
D25N	0.54	30	10	9.0	18	31	98
D27A	0.48	29	11	9.2	18	31	98
D27N	0.52	30	10	9.0	18	31	98
T65A	0.49	29	11	9.4	18	32	99
T65S	0.50	30	10	9.0	18	31	98
R149K	0.50	29	11	9.2	18	31	98
H152F	0.47	28	11	9.3	18	31	98
C187A	0.58	30	10	9.0	18	31	98
D193N	0.37	29	11	9.0	18	30	97
D196N	0.42	29	10	9.2	18	32	98
D199N	0.51	29	11	9.1	18	31	98
D201N	0.50	29	11	9.1	18	31	98
K209E	0.47	29	11	9.2	18	31	98
K209R	0.50	29	11	9.1	18	31	98
D232A	0.55	29	11	9.3	18	32	99
D236A	0.55	29	11	9.2	18	31	98
WT KabB	**0.53**	**30**	**10**	**9.1**	**18**	**31**	**98**
KabB N201D	0.56	30	10	9.0	18	31	98

There is no more than 1% difference between the content of the principal secondary structures, such as the α-helices and antiparallel β-sheets, of the WT and all the mutants. The commonly accepted error of the analysis of the CDNN software is 1%-5%.[77] Since the values obtained are within this margin, it was concluded that the selected amino acid mutations did not notably affect the secondary structure of NtdB and therefore, any observed change in behavior in the enzyme, such as loss or increase in activity, corresponds to the mutation itself. The same can be said about the KabB N201D mutation which made no significant difference in the structure of the enzyme.

CD was also used to compare the secondary structure of KabB and NtdB (**Figure 4.34 and Table 4.10**). Because KabB and NtdB are 55% identical, it was expected that the secondary structure of both enzymes would be similar. This was the case and, until a crystal structure of KabB is revealed, it is proposed that any differences observed

between both enzymes came from their 45% difference in the amino acid sequence which may affect the closed-conformation of the protein or the identity of the native substrate.

5 SUMMARY AND CONCLUSIONS

The objectives of this research were to purify and characterize the activity and substrate selectivity of NtdB, to investigate the role of proposed active-site residues by site-directed mutagenesis, and to initiate an investigation of KabB, an NtdB homolog.

NtdB was successfully purified. The function of NtdB was identified as a K6P phosphatase. NtdB does not have a role similar to trehalose-6-phosphate phosphatase (TPP) since it shows no activity in the presence of phosphorylated analogs of NTD. NtdB showed no activity with G6P or K1P which indicates the importance of the location of the amino and phosphate group for substrate recognition. NtdB exhibited a high level of activity in the presence of K6P, suggesting that it is the native substrate.

Site-directed mutagenesis supported the proposed catalytic mechanism and the importance of the catalytic motif in the core domain of the enzyme was confirmed. Essential amino acids of the cap were identified. The mutations of H152, C187, D193, D199, and D201 changed the activity of the enzyme. Mutations of D199 and D201 changed the substrate specificity of the enzyme.

The kinetic constants of NtdB-catalyzed K6P hydrolysis by wild-type and mutant proteins were obtained. For the WT, K_m and k_{cat} are 93 ± 7 µM, and 32 ± 3 s^{-1}, respectively, while k_{cat}/K_m is (3.4 ± 0.4) x 10^5 $M^{-1}s^{-1}$ under the conditions of the experiments. These numbers identifies NtdB as a K6P phosphatase.

KabB was successfully sub-cloned, expressed and purified. The kinetic constants of KabB with K6P were obtained: K_m, and k_{cat} are 500 ± 50 µM, and 0.42 ± 0.04 s^{-1}, respectively, while k_{cat}/K_m is (8.5 ± 1.0) x 10^2 $M^{-1}s^{-1}$. KabB and NtdB have notably different kinetics constants.

NtdB is part of an operon that was previously reported to produce NTD. Based on the results obtained, the end product of the *ntd* operon is kanosamine, which is created by

a novel biosynthetic pathway. Each enzyme of this pathway requires new Enzyme Commission (EC) numbers because the each catalyzes novel reactions, not observed before. The dephosphorylation of K6P has not been reported before. Since the initial hypothesis was that the operon produces NTD, this would lead to the conclusion that other enzyme(s) are required to achieve this product, enzymes found in *E. coli* and *B. subtilis*, but not in *B. cereus*. The next step would be the investigation of these bacterial species to identify such enzyme(s). The identification of a novel kanosamine pathway will be useful to other researchers investigating the origin of aminoglycoside antibiotics, who can look for this or similar pathways in other organisms.

6 REFERENCES

1. Numata, K.; Yamamoto, H.; Hatori, M.; Miyaki, T.; Kawaguchi, H., Isolation Of An Aminoglycoside Hypersensitive Mutant And Its Application In Screening. *Journal of Antibiotics* **1986,** *39* (7), 994-1000.

2. Tsuno, T.; Ikeda, C.; Numata, K.; Tomita, K.; Konishi, M.; Kawaguchi, H., 3,3'-Neotrehalosadiamine (BMY-28251), A New Aminosugar Antibiotic. *Journal of Antibiotics* **1986,** *39* (7), 1001-1003.

3. Dumitrescu, O.; Dauwalder, O.; Boisset, S.; Reverdy, M.-E.; Tristan, A.; Vandenesch, F., Staphylococcus aureus resistance to antibiotics: key points in 2010. *M S-Medecine Sciences* **2010,** *26* (11), 943-949.

4. Holden, M. T. G.; Lindsay, J. A.; Corton, C.; Quail, M. A.; Cockfield, J. D.; Pathak, S.; Batra, R.; Parkhill, J.; Bentley, S. D.; Edgeworth, J. D., Genome Sequence of a Recently Emerged, Highly Transmissible, Multi-Antibiotic- and Antiseptic-Resistant Variant of Methicillin-Resistant Staphylococcus aureus, Sequence Type 239 (TW). *Journal of Bacteriology* **2010,** *192* (3), 888-892.

5. Bratu, S.; Landman, D.; Haag, R.; Recco, R.; Eramo, A.; Alam, M.; Quale, J., Rapid spread of carbapenem-resistant Klebsiella pneumoniae in New York City - A new threat to our antibiotic armamentarium. *Archives of Internal Medicine* **2005,** *165* (12), 1430-1435.

6. Leavitt, A.; Navon-Venezia, S.; Chmelnitsky, I.; Schwaber, M. J.; Carmeli, Y., Emergence of KPC-2 and KPC-3 in carbapenem-resistant Klebsiella pneumoniae strains in an Israeli hospital. *Antimicrobial Agents and Chemotherapy* **2007,** *51* (8), 3026-3029.

7. Numata, K. I.; Satoh, F.; Hatori, M.; Miyaki, T.; Kawaguchi, H., Isolation Of 3,3'-Neotrehalosadiamine (Bmy-28251) From A Butirosin-Producing Organism. *Journal of Antibiotics* **1986,** *39* (9), 1346-1348.

8. Inaoka, T.; Takahashi, K.; Yada, H.; Yoshida, M.; Ochi, K., RNA Polymerase Mutation Activates The Production Of A Dormant Antibiotic 3,3'-Neotrehalosadiamine Via An Autoinduction Mechanism In Bacillus Subtilis. *Journal of Biological Chemistry* **2004,** *279* (5), 3885-3892.

9. Milner, J. L.; SiloSuh, L.; Lee, J. C.; He, H. Y.; Clardy, J.; Handelsman, J., Production of kanosamine by Bacillus cereus UW85. *Applied and Environmental Microbiology* **1996,** *62* (8), 3061-3065.

10. Janiak, A. M.; Milewski, S., Mechanism of antifungal action of kanosamine. *Medical Mycology* **2001,** *39* (5), 401-408.

11. Iwai, Y.; Tanaka, H.; Oiwa, R.; Shimizu, S.; Omura, S., Studies On Bacterial-Cell Wall Inhibitors .3. 3-Amino-3-Deoxy-D-Glucose, An Inhibitor Of Bacterial-Cell Wall Synthesis. *Biochimica Et Biophysica Acta* **1977,** *498* (1), 223-228.

12. Fusetani, N.; Ejima, D.; Matsunaga, S.; Hashimoto, K.; Itagaki, K.; Akagi, Y.; Taga, N.; Suzuki, K., 3-Amino-3-Deoxy-D-Glucose - An Antibiotic Produced By A Deep-Sea Bacterium. *Experientia* **1987,** *43* (4), 464-465.

13. Dolak, L. A.; Castle, T. M.; Laborde, A. L., 3-Trehalosamine, A New Disaccharide Antibiotic. *Journal of Antibiotics* **1980,** *33* (7), 690-694.

14. Inaoka, T.; Ochi, K., Glucose uptake pathway-specific regulation of synthesis of neotrehalosadiamine, a novel autoinducer produced in Bacillus subtilis. *Journal of Bacteriology* **2007,** *189* (1), 65-75.

15. Langill, D. M. Synthesis of neotrehalose; kinetics and mutagenesis of NtdC. University of Saskatchewan, Saskatoon, SK, 2010.

16. van Straaten, K. E.; Langill, D. M.; Palmer, D. R. J.; Sanders, D. A. R., Purification, crystallization and preliminary X-ray analysis of NtdA, a putative pyridoxal phosphate-dependent aminotransferase from Bacillus subtilis. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **2009,** *65*, 426-429.

17. Vetter, N. D.; Palmer, D. R. J., *Unpublished results*: 2012.

18. van Straaten, K. E.; Langill, D. M.; Palmer, D. R. J.; Sanders, To be published, 2012.

19. Elbein, A. D.; Pan, Y. T.; Pastuszak, I.; Carroll, D., New insights on trehalose: a multifunctional molecule. *Glycobiology* **2003,** *13* (4), 17R-27R.

20. Thompson, J. D.; Higgins, D. G.; Gibson, T. J., Clustal-W - Improving The Sensitivity Of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties And Weight Matrix Choice. *Nucleic Acids Research* **1994,** *22* (22), 4673-4680.

21. Henikoff, S.; Henikoff, J. G., Amino-Acid Substitution Matrices From Protein Blocks. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** *89* (22), 10915-10919.

22. Gouet, P.; Courcelle, E.; Stuart, D. I.; Metoz, F., ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **1999,** *15* (4), 305-308.

23. Kevany, B. M.; Rasko, D. A.; Thomas, M. G., Characterization of the Complete Zwittermicin A Biosynthesis Gene Cluster from Bacillus cereus. *Applied and Environmental Microbiology* **2009,** *75* (4), 1144-1155.

24. Guo, J. T.; Frost, J. W., Kanosamine biosynthesis: A likely source of the aminoshikimate pathway's nitrogen atom. *Journal of the American Chemical Society* **2002,** *124* (36), 10642-10643.

25. Engels, S. Sugars in early and late polyketide biosynthesis: Functional studies of *rifL*, *rifK* and *rifM* in rifamycin biosynthesis; Towards the characterisation of a PKS gene cluster from *Streptomyces* sp. GW2/5831, encoding the biosynthesis of the polycyclic xanthone IB-00208. University of Bonn, University of Bonn, 2009.

26. Arakawa, K.; Muller, R.; Mahmud, T.; Yu, T. W.; Floss, H. G., Characterization of the early stage aminoshikimate pathway in the formation of 3-amino-5-hydroxybenzoic acid: The RifN protein specifically converts kanosamine into kanosamine 6-phosphate. *Journal of the American Chemical Society* **2002,** *124* (36), 10644-10645.

27. Westheimer, F. H., Why Nature Chose Phosphates. *Science* **1987,** *235* (4793), 1173-1178.

28. Florian, J.; Warshel, A., Phosphate ester hydrolysis in aqueous solution: Associative versus dissociative mechanisms. *Journal of Physical Chemistry B* **1998,** *102* (4), 719-734.

29. Florian, J.; Warshel, A., A fundamental assumption about OH- attack in phosphate ester hydrolysis is not fully justified. *Journal of the American Chemical Society* **1997,** *119* (23), 5473-5474.

30. Lad, C.; Williams, N. H.; Wolfenden, R., The rate of hydrolysis of phosphomonoester dianions and the exceptional catalytic proficiencies of protein and inositol phosphatases. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** *100* (10), 5607-5610.

31. Wolfenden, R.; Snider, M. J., The depth of chemical time and the power of enzymes as catalysts. *Accounts of Chemical Research* **2001,** *34* (12), 938-945.

32. Rusnak, F.; Mertz, P., Calcineurin: Form and function. *Physiological Reviews* **2000,** *80* (4), 1483-1521.

33. Schenk, G.; Boutchard, C. L.; Carrington, L. E.; Noble, C. J.; Moubaraki, B.; Murray, K. S.; de Jersey, J.; Hanson, G. R.; Hamilton, S., A purple acid phosphatase from sweet potato contains an antiferromagnetically coupled binuclear Fe-Mn center. *Journal of Biological Chemistry* **2001,** *276* (22), 19084-19088.

34. Egloff, M. P.; Cohen, P. T. W.; Reinemer, P.; Barford, D., Crystal-Structure Of The Catalytic Subunit Of Human Protein Phosphatase-1 And Its Complex With Tungstate. *Journal of Molecular Biology* **1995,** *254* (5), 942-959.

35. Shi, Y., Serine/Threonine Phosphatases: Mechanism through Structure. *Cell* **2009,** *139* (3), 468-484.

36. Changela, A.; Ho, C. K.; Martins, A.; Shuman, S.; Mondragon, A., Structure and mechanism of the RNA triphosphatase component of mammalian mRNA capping enzyme. *Embo Journal* **2001,** *20* (10), 2575-2586.

37. McCain, D. F.; Catrina, I. E.; Hengge, A. C.; Zhang, Z. Y., The catalytic mechanism of Cdc25A phosphatase. *Journal of Biological Chemistry* **2002,** *277* (13), 11190-11200.

38. Baker, A. S.; Ciocci, M. J.; Metcalf, W. W.; Kim, J.; Babbitt, P. C.; Wanner, B. L.; Martin, B. M.; Dunaway-Mariano, D., Insights into the mechanism of catalysis by the P-C bond-cleaving enzyme phosphonoacetaldehyde hydrolase derived from gene sequence analysis and mutagenesis. *Biochemistry* **1998,** *37* (26), 9305-9315.

39. Wang, W. Q.; Sun, J. P.; Zhang, Z. Y., An overview of the protein tyrosine phosphatase superfamily. *Current Topics in Medicinal Chemistry* **2003,** *3* (7), 739-748.

40. Stec, B.; Holtz, K. M.; Kantrowitz, E. R., A revised mechanism for the alkaline phosphatase reaction involving three metal ions. *Journal of Molecular Biology* **2000,** *299* (5), 1303-1311.

41. de Backer, M.; McSweeney, S.; Rasmussen, H. B.; Riise, B. W.; Lindley, P.; Hough, E., The 1.9 angstrom crystal structure of heat-labile shrimp alkaline phosphatase. *Journal of Molecular Biology* **2002,** *318* (5), 1265-1274.

42. Ostanin, K.; Harms, E. H.; Stevis, P. E.; Kuciel, R.; Zhou, M. M.; Vanetten, R. L., Overexpression, Site-Directed Mutagenesis, And Mechanism Of Escherichia-Coli Acid-Phosphatase. *Journal of Biological Chemistry* **1992,** *267* (32), 22830-22836.

43. Ghosh, A.; Shieh, J. J.; Pan, C. J.; Sun, M. S.; Chou, J. Y., The catalytic center of glucose-6-phosphatase - HIS176 is the nucleophile forming the phosphohistidine-enzyme intermediate during catalysis. *Journal of Biological Chemistry* **2002,** *277* (36), 32837-32842.

44. Burroughs, A. M.; Allen, K. N.; Dunaway-Mariano, D.; Aravind, L., Evolutionary genomics of the HAD superfamily: Understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *Journal of Molecular Biology* **2006,** *361* (5), 1003-1034.

45. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J. H.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **1997,** *25* (17), 3389-3402.

46. Koonin, E. V.; Tatusov, R. L., Computer-Analysis Of Bacterial Haloacid Dehalogenases Defines A Large Superfamily Of Hydrolases With Diverse Specificity - Application Of An Iterative Approach To Database Search. *Journal of Molecular Biology* **1994,** *244* (1), 125-132.

47. Nocek, B.; Stein, A.; Wu, R.; Jedrzejczak, R.; Joachimiak, A., Crystal structure of yhjK (haloacid dehalogenase-like hydrolase protein) from Bacillus subtilis. 2009.

48. Rao, S. T.; Rossmann, M. G., Comparison Of Super-Secondary Structures In Proteins. *Journal of Molecular Biology* **1973,** *76* (2), 241-&.

49. Kleiger, G.; Eisenberg, D., GXXXG and GXXXA motifs stabilize FAD and NAD(P)-binding Rossmann folds through C-alpha-H center dot center dot center dot O hydrogen bonds and van der Waals interactions. *Journal of Molecular Biology* **2002,** *323* (1), 69-76.

50. Brakoulias, A.; Jackson, R. M., Towards a structural classification of phosphate binding sites in protein-nucleotide complexes: An automated all-against-all structural comparison using geometric matching. *Proteins-Structure Function and Bioinformatics* **2004,** *56* (2), 250-260.

51. Fieulaine, S.; Lunn, J. E.; Borel, F.; Ferrer, J. L., The structure of a cyanobacterial sucrose-phosphatase reveals the sugar tongs that release free sucrose in the cell. *Plant Cell* **2005,** *17* (7), 2049-2058.

52. Regni, C.; Tipton, P. A.; Beamer, L. J., Crystal structure of PMM/PGM: An enzyme in the biosynthetic pathway of P-aeruginosa virulence factors. *Structure* **2002,** *10* (2), 269-279.

53. Rao, K. N.; Kumaran, D.; Seetharaman, J.; Bonanno, J. B.; Burley, S. K.; Swaminathan, S., Crystal structure of trehalose-6-phosphate phosphatase-related protein: Biochemical and biological implications. *Protein Science* **2006,** *15* (7), 1735-1744.

54. Kawamura, T.; Watanabe, N.; Tanaka, I., Structure of mannosyl-3-phosphoglycerate phosphatase from Pyrococcus horikoshii. *Acta Crystallographica Section D-Biological Crystallography* **2008,** *64*, 1267-1276.

55. Aravind, L.; Galperin, M. Y.; Koonin, E. V., The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. *Trends in Biochemical Sciences* **1998,** *23* (4), 127-129.

56. Seal, S. N.; Rose, Z. B., Characterization Of A Phosphoenzyme Intermediate In The Reaction Of Phosphoglycolate Phosphatase. *Journal of Biological Chemistry* **1987,** *262* (28), 13496-13500.

57. Collet, J. F.; Gerin, I.; Rider, M. H.; VeigadaCunha, M.; VanSchaftingen, E., Human L-3-phosphoserine phosphatase: Sequence, expression and evidence for a phosphoenzyme intermediate. *Febs Letters* **1997,** *408* (3), 281-284.

58. Collet, J. F.; Stroobant, V.; Van Schaftingen, E., Mechanistic studies of phosphoserine phosphatase, an enzyme related to P-type ATPases. *Journal of Biological Chemistry* **1999,** *274* (48), 33985-33990.

59. Cronin, A.; Homburg, S.; Duerk, H.; Richter, I.; Adamskal, M.; Frere, F.; Arand, M., Insights into the Catalytic Mechanism of Human sEH Phosphatase by Site-Directed Mutagenesis and LC-MS/MS Analysis. *Journal of Molecular Biology* **2008,** *383* (3), 627-640.

60. Lahiri, S. D.; Zhang, G. F.; Dunaway-Mariano, D.; Allen, K. N., Caught in the act: The structure of phosphorylated beta-phosphoglucomutase from Lactococcus lactis. *Biochemistry* **2002,** *41* (26), 8351-8359.

61. Morais, M. C.; Zhang, W. H.; Baker, A. S.; Zhang, G. F.; Dunaway-Mariano, D.; Allen, K. N., The crystal structure of Bacillus cereus phosphonoacetaldehyde hydrolase: Insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. *Biochemistry* **2000,** *39* (34), 10385-10396.

62. Wang, W. R.; Kim, R.; Jancarik, J.; Yokota, H.; Kim, S. H., Crystal structure of phosphoserine phosphatase from Methanococcus jannaschii, a hyperthermophile, at 1.8 angstrom resolution. *Structure* **2001,** *9* (1), 65-71.

63. Rinaldo-Matthis, A.; Rampazzo, C.; Reichard, P.; Bianchi, V.; Nordlund, P., Crystal structure of a human mitochondrial deoxyribonucleotidase. *Nature Structural Biology* **2002,** *9* (10), 779-787.

64. Kim, Y. C.; Yakunin, A. F.; Kuznetsova, E.; Xu, X. H.; Pennycooke, M.; Gu, J.; Cheung, F.; Proudfoot, M.; Arrowsmith, C. H.; Joachimiak, A.; Edwards, A. M.; Christendat, D., Structure- and function-based characterization of a new phosphoglycolate phosphatase from Thermoplasma acidophilum. *Journal of Biological Chemistry* **2004,** *279* (1), 517-526.

65. Lu, Z. B.; Dunaway-Mariano, D.; Allen, K. N., HAD superfarnily phosphotransferase substrate diversification: Structure and function analysis of HAD subclass IIB sugar phosphatase BT4131. *Biochemistry* **2005,** *44* (24), 8684-8696.

66. Lahiri, S. D.; Zhang, G. F.; Dai, J. Y.; Dunaway-Mariano, D.; Allen, K. N., Analysis of the substrate specificity loop of the HAD superfamily cap domain. *Biochemistry* **2004,** *43* (10), 2812-2820.

67. Omene, E.; Vetter, N.; Palmer, D. R. J., Unpublished results: 2010.

68. Benson, D. A.; Cavanaugh, M.; Clark, K.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Sayers, E. W., GenBank. *Nucleic Acids Research* **2013,** *41* (D1), D36-D42.

69. Kibbe, W. A., OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Research* **2007,** *35*, W43-W46.

70. Vincze, T.; Posfai, J.; Roberts, R. J., NEBcutter: a program to cleave DNA with restriction enzymes. *Nucleic Acids Research* **2003,** *31* (13), 3688-3691.

71. Guex, N.; Peitsch, M. C., SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **1997,** *18* (15), 2714-2723.

72. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T., The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **2006,** *22* (2), 195-201.

73. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C., SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* **2003,** *31* (13), 3381-3385.

74. Wishart, D. S.; Knox, C.; Guo, A. C.; Eisner, R.; Young, N.; Gautam, B.; Hau, D. D.; Psychogios, N.; Dong, E.; Bouatra, S.; Mandal, R.; Sinelnikov, I.; Xia, J.; Jia, L.; Cruz, J. A.; Lim, E.; Sobsey, C. A.; Shrivastava, S.; Huang, P.; Liu, P.; Fang, L.; Peng, J.; Fradette, R.; Cheng, D.; Tzur, D.; Clements, M.; Lewis, A.; De Souza, A.; Zuniga, A.; Dawe, M.; Xiong, Y.; Clive, D.; Greiner, R.; Nazyrova, A.; Shaykhutdinov, R.; Li, L.; Vogel, H. J.; Forsythe, I., HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research* **2009,** *37*, D603-D610.

75. Sambrook, J.; Russell, D. W., *Molecular cloning: A laboratory manual*. 2001.

76. Johnson, W. C., Protein Secondary Structure And Circular-Dichroism - A Practical Guide. *Proteins-Structure Function and Genetics* **1990,** *7* (3), 205-214.

77. Bohm, G.; Muhr, R.; Jaenicke, R., Quantitative-Analysis Of Protein Far Uv Circular-Dichroism Spectra By Neural Networks. *Protein Engineering* **1992,** *5* (3), 191-195.

78. Parsons, J. F.; Lim, K.; Tempczyk, A.; Krajewski, W.; Eisenstein, E.; Herzberg, O., From structure to function: YrbI from Haemophilus influenzae (HI1679) is a phosphatase. *Proteins-Structure Function and Genetics* **2002,** *46* (4), 393-404.

79. Harder, K. W.; Owen, P.; Wong, L. K. H.; Aebersold, R.; Clarklewis, I.; Jirik, F. R., Characterization And Kinetic-Analysis Of The Intracellular Domain Of Human-Protein-Tyrosine-Phosphatase-Beta (HPTP-Beta) Using Synthetic Phosphopeptides. *Biochemical Journal* **1994,** *298*, 395-401.

80. Webb, M. R., A Continuous Spectrophotometric Assay For Inorganic-Phosphate And For Measuring Phosphate Release Kinetics In Biological-Systems. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** *89* (11), 4884-4887.

81. Weber, K.; Osborn, M., Reliability Of Molecular Weight Determinations By Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis. *Journal of Biological Chemistry* **1969,** *244* (16), 4406-&.

82. Dunker, A. K.; Rueckert, R. R., Observations On Molecular Weight Determinations On Polyacrylamide Gel. *Journal of Biological Chemistry* **1969,** *244* (18), 5074-&.

83. Hirel, P. H.; Schmitter, J. M.; Dessen, P.; Fayat, G.; Blanquet, S., Extent Of N-Terminal Methionine Excision From Escherichia-Coli Proteins Is Governed By The Side-Chain Length Of The Penultimate Amino-Acid. *Proceedings of the National Academy of Sciences of the United States of America* **1989,** *86* (21), 8247-8251.

84. Tramer, J., Simple And Rapid Methods For The Estimation Of Bacterial Phosphatases Using Di-Sodium Para-Nitrophenylphosphate As Substrate. *Journal of Dairy Research* **1952,** *19* (3), 275-287.

85. Shin, D. H.; Roberts, A.; Jancarik, J.; Yokota, H.; Kim, R.; Wemmer, D. E.; Kim, S. H., Crystal structure of a phosphatase with a unique substrate binding domain from Thermotoga maritima. *Protein Science* **2003,** *12* (7), 1464-1472.

86. Soyenkoff, B., A Micromethod Of Phosphate Determination. *Journal of Biological Chemistry* **1947,** *168* (2), 447-457.

87. Baykov, A. A.; Evtushenko, O. A.; Avaeva, S. M., A Malachite Green Procedure For Ortho-Phosphate Determination And Its Use In Alkaline Phosphatase-Based Enzyme-Immunoassay. *Analytical Biochemistry* **1988,** *171* (2), 266-270.

88. Kormish, J. D.; McGhee, J. D., The C-elegans lethal gut-obstructed gob-1 gene is trehalose-6-phosphate phosphatase. *Developmental Biology* **2005,** *287* (1), 35-47.

89. Nelson, D. L.; Cox, M. M., *Lehninger'a Principles of Biochemistry*. 4th Edition ed.; Freeman & Co.: WH, 2004.

90. Khersonsky, O.; Tawfik, D. S., Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective. In *Annual Review of Biochemistry, Vol 79*, Kornberg, R. D.; Raetz, C. R. H.; Rothman, J. E.; Thorner, J. W., Eds. 2010; Vol. 79, pp 471-505.

91. Glasner, M. E.; Gerlt, J. A.; Babbitt, P. C., Evolution of enzyme superfamilies. *Current Opinion in Chemical Biology* **2006,** *10* (5), 492-497.

92. Beassoni, P. R.; Otero, L. H.; Massimelli, M. J.; Lisa, A. T.; Domenech, C. E., Critical active-site residues identified by site-directed mutagenesis in Pseudomonas aeruginosa phosphorylcholine phosphatase, a new member of the haloacid dehalogenases hydrolase superfamily. *Current Microbiology* **2006,** *53* (6), 534-539.

93. Selengut, J. D., MDP-1 is a new and distinct member of the haloacid dehalogenase family of aspartate-dependent phosphohydrolases. *Biochemistry* **2001,** *40* (42), 12704-12711.

94. Kurihara, T.; Liu, J. Q.; Nardidei, V.; Koshikawa, H.; Esaki, N.; Soda, K., Comprehensive Site-Directed Mutagenesis Of L-2-Halo Acid Dehalogenase To Probe Catalytic Amino-Acid-Residues. *Journal of Biochemistry* **1995,** *117* (6), 1317-1322.

95. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **2004,** *60*, 2126-2132.

96. Barnes, M. R.; Gray, I. C., *Bioinformatics for geneticists*. 2003; p xiv + 408 pp.-xiv + 408 pp.

97. Bhattacharyya, R.; Pal, D.; Chakrabarti, P., Disulfide bonds, their stereospecific environment and conservation in protein structures. *Protein Engineering Design & Selection* **2004,** *17* (11), 795-808.

98. Kokke, W., Determination Of Optical Purity Using Circular-Polarization Of Luminescence. *Journal of the American Chemical Society* **1974,** *96* (8), 2627-2628.

99. Greenfield, N. J., Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **2006,** *1* (6), 2876-2890.

100. Johnson, W. C., Secondary Structure Of Proteins Through Circular-Dichroism Spectroscopy. *Annual Review of Biophysics and Biophysical Chemistry* **1988,** *17*, 145-166.

101. Holzwart.G; Doty, P., Ultraviolet Circular Dichroism Of Polypeptides. *Journal of the American Chemical Society* **1965,** *87* (2), 218-&.

102. Greenfie.N; Fasman, G. D., Computed Circular Dichroism Spectra For Evaluation Of Protein Conformation. *Biochemistry* **1969,** *8* (10), 4108-&.

103. Venyaminov, S. Y.; Baikalov, I. A.; Shen, Z. M.; Wu, C. S. C.; Yang, J. T., Circular Dichroic Analysis Of Denatured Proteins - Inclusion Of Denatured Proteins In The Reference Set. *Analytical Biochemistry* **1993,** *214* (1), 17-24.